UNIVERSAL REGULAR CONTROL FOR GENERIC SEMILINEAR SYSTEMS

JAIRO BOCHI AND NICOLAS GOURMELON

Abstract. We consider discrete-time projective semilinear control systems $\xi_{t+1} = A(u_t) \cdot \xi_t$, where the states ξ_t are in projective space $\mathbb{RP}^{d-1}$, inputs u_t are in a manifold $\mathcal{U}$ of arbitrary dimension, and $A \colon \mathcal{U} \to \mathrm{GL}(d, \mathbb{R})$ is a fixed differentiable map.

An input sequence $(u_0, \dots, u_{N-1})$ is called universally regular if for any initial state $\xi_0 \in \mathbb{RP}^{d-1}$, the derivative of the time-N state with respect to the inputs is onto.

In this paper we deal with the universal regularity of constant input sequences $(u_0, \dots, u_0)$. Our main result states that for generic such control systems, all constant inputs of sufficient length N are universally regular, except for a discrete set. More precisely, the conclusion holds for a C^2-open and C^∞-dense set of maps A. We also show that the inputs on that discrete set are nearly universally regular; indeed there is a unique non-regular initial state, and its corank is 1.

In order to establish the result, we study the spaces of bilinear control systems. We show that the codimension of the set of systems for which the zero input is not universally regular coincides with the dimension of the control space. The proof is based on careful matrix analysis and some elementary algebraic geometry. Then the main result follows by applying standard transversality theorems.

1. Introduction

1.1. **Basic definitions and some questions.** Consider discrete-time control systems of the form:

$$x_{t+1} = F(x_t, u_t), \qquad (t = 0, 1, 2, \dots) \tag{1.1}$$

where $F \colon \mathcal{X} \times \mathcal{U} \to \mathcal{X}$ is any map. We will always assume that the space $\mathcal{X}$ of states and the space $\mathcal{U}$ of controls are manifolds, and that the map F is continuously differentiable.

A sequence $(x_0, \dots, x_N; u_0, \dots, u_{N-1})$ satisfying (1.1) is called a trajectory of length N; it is uniquely determined by the initial state x_0 and the input $(u_0, \dots, u_{N-1})$. Let ϕ_N denote the time-N transition map, which gives the final state as a function of the initial state and the input:

$$x_N = \phi_N(x_0; u_0, \dots, u_{N-1}). \tag{1.2}$$

We say that the system (1.1) is *accessible* from x_0 in time N if the set $\phi_N(\{x_0\} \times \mathcal{U}^N)$ of final states that can be reached from the initial state x_0 has nonempty interior.

Date: January, 2012.

2010 *Mathematics Subject Classification.* 93C10; 93B05, 93C55.

Key words and phrases. Discrete-time systems; semilinear systems; bilinear systems; universal regular control.

J.B. was partially supported by CNPq, Université Bordeaux I, and Rede Franco–Brasileira em Matemática. N.G. was partially supported by FAPERJ.

The implicit function theorem gives a sufficient condition for accessibility. If the derivative of the map $\phi_N(x_0;\cdot)$ at input $(u_0,\dots,u_{N-1})$ is an onto linear map[1] then we say that the trajectory determined by $(x_0;u_0,\dots,u_{N-1})$ is *regular*. So the existence of such a regular trajectory implies that the system is accessible from x_0 in time N.

Let us call an input $(u_0,\dots,u_{N-1})$ *universally regular* if for every $x_0 \in \mathcal{X}$, the trajectory determined by $(x_0;u_0,\dots,u_{N-1})$ is regular; otherwise the input is called singular.

This concept is central in the present paper; it was introduced by Sontag[2] in [So] in the context of continuous-time control systems. The discrete-time analogue was considered by Sontag and Wirth in [SW]. They showed that if the system (1.1) is accessible from every initial condition x_0 in uniform time N then universally regular inputs do exist, provided one assumes the map F to be analytic. In fact, under those hypotheses they showed that universally regular inputs are abundant: in the space of inputs of sufficiently large length, those that are not universally regular form a set of positive codimension.

In this paper, we are interested in control systems (1.1) where the next state x_{t+1} depends linearly on the previous state x_t (but non-linearly on u_t, in general). This means that the state space is $\mathbb{K}^d$, where $d \geqslant 2$ and $\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$ and that (1.1) now takes the form:

$$x_{t+1} = A(u_t)\cdot x_t, \qquad \text{where } A\colon \mathcal{U} \to \mathrm{Mat}_{d\times d}(\mathbb{K}). \tag{1.3}$$

Following [CK1], we call this a *semilinear control system*.

In the case that the map A above takes values in the set $\mathrm{GL}(d,\mathbb{K})$ of invertible matrices, we consider the corresponding projectivized control system:

$$\xi_{t+1} = A(u_t)\cdot \xi_t, \tag{1.4}$$

where the states ξ_t take value in the projective space $\mathbb{K}\mathrm{P}^{d-1} = \mathbb{K}^d_*/\mathbb{K}_*$. We call this a *projective semilinear control system*. The projectivized system is also a useful tool for the study of the original system (1.3): see e.g. [Wi, CK2].

Universally regular inputs for projective semilinear control systems were first considered by Wirth in [Wi]. Under his working hypotheses, the existence and abundance of such inputs is guaranteed by the aforementioned result of [SW]; then he uses universally regular inputs to obtain global controllability properties.

The purpose of this paper is to establish results on the existence and abundance of universally regular inputs for projective semilinear control systems. Differently from [SW, Wi], we will not necessarily assume our systems to be analytic. Let us consider systems (1.4) with $\mathbb{K}=\mathbb{R}$ and $A\colon \mathcal{U}\to \mathrm{GL}(d,\mathbb{R})$ a map of some class C^r, with $r \geqslant 1$. To compensate for less rigidity, we do not try to obtain results that work for all C^r maps A, but only for *generic* ones, i.e., those maps in a residual[3] subset, or, even better, in an open dense subset.

To make things more precise, assume $\mathcal{U}$ is a C^∞ (real) manifold without boundary.[4] We will always consider the space $C^r(\mathcal{U},\mathrm{GL}(d,\mathbb{R}))$ endowed with the strong C^r topology[5].

[1] This condition is usually written in terms of the rank of a certain matrix and it is usually called the *rank condition*.

[2] Sontag calls these inputs universally nonsingular; we follow the terminology of [Wi, CK2].

[3] Recall that a subset of a Baire space is called *residual* if it is a countable intersection of open dense subsets.

[4] Moreover, all manifolds are assumed to be Hausdorff paracompact with a countable base of open sets, and of finite dimension.

[5] See e.g. [Hi]. Note that in the case that $\mathcal{U}$ is compact, this coincides with the usual uniform C^r topology.

Hence the first question we pose is this:

> Taking N sufficiently large, is it true that for C^r-generic maps A, the set of universally regular inputs in $\mathcal{U}^N$ is itself generic?

It turns out that this question has a positive answer. Actually, we show in [BG3] that for r great enough, for maps A in a C^r open and dense set, all inputs in $\mathcal{U}^N$ are universally regular, except for those in a stratified closed set of positive codimension. So another natural question is this:

> Fixed parameters d, $\dim \mathcal{U}$, N, and r, what is the minimum codimension of the set of singular inputs in $\mathcal{U}^N$ that can occur for C^r-generic maps $A \colon \mathcal{U} \to \mathrm{GL}(d, \mathbb{R})$?

This question seems to be very difficult. However, we do have a sharp estimate if we restrict ourselves to the subset of $\mathcal{U}^N$ formed by *non-resonant inputs*, namely those inputs $(u_0, \dots, u_{N-1})$ such that $u_i \neq u_j$ whenever $i \neq j$. To investigate what happens for resonant inputs is a much tougher job.

In this paper we consider the most resonant case. Define a *constant* input of length N as an element of $\mathcal{U}^N$ of the form $(u_0, u_0, \dots, u_0)$. We propose ourselves to study universal regularity of inputs of this form. A possible interpretation is this: Suppose the system is controlled by a "lever" that is very hard to move. Then we want to know what positions of the lever are universally regular. For those positions it is possible to perturb the state of the system in any desired direction with only small moves on the lever.

1.2. **The main result.**

Our main result says that generically the singular constant inputs form a very small set:

Theorem 1.1. *Given $d \geqslant 2$ and $m \geqslant 1$, there exists $N \geqslant 1$ with the following properties. Let $\mathcal{U}$ be a smooth m-dimensional manifold without boundary. Then there exists a C^2-open C^∞-dense subset $\mathcal{O}$ of $C^2(\mathcal{U}, \mathrm{GL}(d, \mathbb{R}))$ such that for every system* (1.4) *with $A \in \mathcal{O}$, all constant inputs of length N are universally regular, except for those in a zero-dimensional (i.e., discrete) set.*

By saying that a subset $\mathcal{O}$ of $C^2(\mathcal{U}, \mathrm{GL}(d, \mathbb{R}))$ is C^∞-dense, we mean that for all $r \geqslant 2$, the intersection of $\mathcal{O}$ with $C^r(\mathcal{U}, \mathrm{GL}(d, \mathbb{R}))$ is dense in $C^r(\mathcal{U}, \mathrm{GL}(d, \mathbb{R}))$.

It is remarkable that the generic dimension of the set of singular constant inputs (namely, 0) does not depend on the dimension m of the control space $\mathcal{U}$, neither on the dimension $d-1$ of the state space. A partial explanation for this phenomenon is the following: First, the obstruction to universal regularity of the input $(u, u, \dots, u)$ is the combined degeneracy of the matrix $A(u)$ and of the derivatives of A at u. If m is small then the image of the generic map A will avoid too degenerate matrices, which increases the chances of obtaining universal regularity. If m is large then more degenerate matrices $A(u)$ will inevitably appear; however the large number of control parameters compensates, so universal control is still likely.

The singular inputs that appear in Theorem 1.1 are not only rare; we also show that they are "almost" universally regular:

Theorem 1.2 (Addendum to Theorem 1.1)**.** *The set $\mathcal{O} \subset C^2(\mathcal{U}, \mathrm{GL}(d, \mathbb{R})$ in Theorem 1.1 can be taken with the following additional properties: If $A \in \mathcal{O}$ and a constant input $(u, \dots, u)$ of length N is singular then:*

1. *There is a single direction $\xi_0 \in \mathbb{RP}^{d-1}$ such the the corresponding trajectory of system* (1.4) *is not regular.*
2. *The derivative of the map $\phi_N(\xi_0; \cdot)$ at input $(u, \dots, u)$ has corank*[6] 1.

[6]The *corank* of a linear map $L \colon V \to W$ is the number $\dim W - \dim L(V)$.

To sum up, for generic systems (1.4), the universal regularity of constant inputs can fail only in the weakest possible way: there is at most one non-regular state, which can be moved in all directions but one.

Remark 1.3. We actually have a very precise description of the singular inputs that appear in Theorem 1.2. We show that these singular inputs can be unremovable by perturbations, and therefore Theorem 1.1 is optimal in the sense that there are C^2-open (actually even C^1-open) sets of maps A for which the set of singular constant inputs is nonempty. Also, by C^1-perturbing any A in those C^2-open sets, one can obtain an infinite number of singular constant inputs. In particular, it is not possible to choose $\mathcal{O}$ to be C^1-open in the statement of the Theorem 1.1. See Appendix B.

Remark 1.4. The integer N is a function of d and m we did not try to estimate precisely. However, we know that it is at most d^2 (see Remark 1.7).

Remark 1.5. In the case of complex matrices (i.e., $\mathbb{K} = \mathbb{C}$), we have a corresponding version of Theorem 1.1 where the maps A are analytic; see Appendix C.

1.3. **Reduction to the study of the set of poor data.** The bulk of the proof of Theorem 1.1 consists on the computation of the dimension of certain canonical sets, as we now explain.

We fix $A\colon \mathcal{U} \to \mathrm{GL}(d,\mathbb{K})$ and consider the projective semilinear system (1.4). Recall that 1-*jet* of the map A at a point $u \in \mathcal{U}$ consists of the first order Taylor approximation of A around u. By the chain rule, the universal regularity of an input $(u_0, u_1, \dots, u_{N-1})$ depends only on the 1-jets of A at points $u_0, \dots, u_{N-1}$.

Let us discuss the case of constant inputs $(u_0, \dots, u_0)$. If we take local coordinates such that $u_0 = 0$ and replace the matrix map $A\colon \mathcal{U} \to \mathrm{GL}(d,\mathbb{K})$ by its linear approximation, system (1.4) becomes:

$$\xi_{t+1} = \left(A + \sum_{j=1}^{m} u_{t,i} C_j\right) \xi_t\,, \quad (t = 0, 1, 2, \dots), \tag{1.5}$$

where $A = A(u_0)$ and $C_1, \dots, C_m$ are the partial derivatives at 0. This is the projectivization of a *bilinear control system* (see [El]). For these systems, the zero input is a distinguished one and the focus of more attention.

To study system (1.5) it is actually more convenient to consider *normalized derivatives* $B_j = C_j A^{-1}$, which intrinsically take values in the Lie algebra $\mathfrak{gl}(d,\mathbb{K})$. Consider the matrix data $\mathbf{A} = (A, B_1, \dots, B_m)$. We will explain how the universal regularity of the zero input is expressed in linear algebraic terms. Recall that the *adjoint operator* of A acts on $\mathfrak{gl}(d,\mathbb{K})$ by the formula $\mathrm{Ad}_A(B) = ABA^{-1}$. Consider the linear subspace $\Lambda_N(\mathbf{A})$ of $\mathfrak{gl}(d,\mathbb{K})$ spanned by the matrices

$$\mathrm{Id} \quad \text{and} \quad (\mathrm{Ad}_A)^i(B_j), \quad (i = 0, \dots, n-1,\ j = 1, \dots, m).$$

(The identity matrix appears because of the projectivization.) Then:

Proposition 1.6. *The constant input* $(0, \dots, 0)$ *of length* N *is universally regular for system* (1.5) *if and only if the space* $\Lambda_N(\mathbf{A})$ *acts transitively on the set* $\mathbb{K}^d_*$ *of nonzero vectors.*

If $\Lambda_N(\mathbf{A})$ acts transitively on $\mathbb{K}^d_*$ for some N, then the data $\mathbf{A}$ is called *rich*; otherwise it is called *poor*.

Remark 1.7. The spaces $\Lambda_N(\mathbf{A})$ form a nested sequence, which thus stabilize after finitely many steps. It is actually easy to see that stabilization occurs at most at time $N = d^2$. Therefore there are two possibilities: either the zero input of length d^2 is universally regular, or the zero inputs of all lengths are singular.[7]

[7]In other words, if you're old enough and still poor then you'll never get rich.

Let $\mathcal{P}_m^{(\mathbb{K})}$ denote the set of poor data.[8] A major part of our work is to study these sets. We prove:

Theorem 1.8. *The set $\mathcal{P}_m^{(\mathbb{R})}$ is closed and semialgebraic, and its codimension in $\mathrm{GL}(d,\mathbb{R}) \times (\mathfrak{gl}(d,\mathbb{R}))^m$ is m.*

Theorem 1.9. *The set $\mathcal{P}_m^{(\mathbb{C})}$ is algebraic, and its (complex) codimension in $\mathrm{GL}(d,\mathbb{C}) \times (\mathfrak{gl}(d,\mathbb{C}))^m$ is m.*

So Theorems 1.8 and 1.9 say how frequent universal regularity of the zero input is in the space of projective bilinear control systems (1.5)

1.4. **Overview of the proofs.** Theorem 1.1 follow rather directly from Theorem 1.8 by applying standard results from transversality theory. More precisely, the fact that the set $\mathcal{P}_m^{(m)}$ is semialgebraic implies that it has a canonical stratification. This permits us to apply Thom's jet transversality theorem and obtain Theorem 1.1.

On the other hand, Theorem 1.8 follows from its complex version Theorem 1.9 by simple abstract arguments.

Thus everything is based on Theorem 1.9. One part of the result is easily obtained: we give examples of small disks of codimension m formed by poor data, so concluding that the codimension of $\mathcal{P}_m^{(\mathbb{C})}$ is at most m.

To prove the other inequality, one could try to exhibit an explicit codimension m set containing all poor data. For $m = 1$ this task is feasible (and we actually perform it, because with these conditions we can actually check universal regularity in concrete examples). However, for $m = 2$ already the task would be very laborious, and to expect to find a general solution seems unrealistic.

Our actual approach to prove the lower bound on the codimension of $\mathcal{P}_m^{(\mathbb{C})}$ is indirect. Crudely speaking, after careful matrix computations, we find some sets in the complement of $\mathcal{P}_m^{(\mathbb{C})}$ that are reasonably "large" (basically in terms of dimension). Then, by using some abstract results of algebraic geometry, we are able to show that $\mathcal{P}_m^{(\mathbb{C})}$ is "small", thus proving the other half of Theorem 1.9.

Let us give more detail about this strategy. We decompose the set $\mathcal{P}_m = \mathcal{P}_m^{(\mathbb{C})}$ into fibers:

$$\mathcal{P}_m = \bigcup_{A \in \mathrm{GL}(d,\mathbb{C})} \{A\} \times \mathcal{P}_m(A), \qquad \mathcal{P}_m(A) \subset [\mathfrak{gl}(d,\mathbb{C})]^m.$$

It is not very difficult to show that for generic A in $\mathrm{GL}(d,\mathbb{C})$, the fiber $\mathcal{P}_m(A)$ has precisely the wanted codimension m. However, for degenerate matrices A, the fiber $\mathcal{P}_m(A)$ may be much bigger. (For example, one can show that if A is an homothecy and $m \leqslant 2d-3$ then $\mathcal{P}_m(A)$ is the whole $[\mathfrak{gl}(d,\mathbb{C})]^m$.) In order to show that $\operatorname{codim} \mathcal{P}_m \geqslant m$, we need to make sure that those degenerate matrices with do not form a large set. More precisely, we show that:

$$\forall k \in \{0,\dots,m\}, \ \operatorname{codim}\big\{A \in \mathrm{GL}(d,\mathbb{C});\ \operatorname{codim} \mathcal{P}_m(A) \leqslant m-k\big\} \geqslant k. \tag{1.6}$$

Let us explain how we prove (1.6). In order to estimate the dimension of $\mathcal{P}_m(A)$ for any matrix $A \in \mathrm{GL}(d,\mathbb{C})$, we consider a quantity $r = r(A)$ which is the least number such that a rich data of the form $(A, C_1, \dots, C_r)$ exists. In particular, if $r = r(A) \leqslant m$ then the following affine space

$$\big\{(C_1, C_2, \dots, C_r, B_{r+1}, \dots, B_m);\ B_j \in \mathfrak{gl}(d,\mathbb{C})\big\} \tag{1.7}$$

is contained in the complement of $\mathcal{P}_m(A)$.

[8] A more precise notation would be $\mathcal{P}_{m,d}^{(\mathbb{K})}$. However, we can think d as fixed; on the other hand it is sometimes useful to change m.

In certain situations, if two algebraic subsets have large enough dimensions then they necessarily intersect; for example, two algebraic curves in the complex projective plane $\mathbb{C}\mathrm{P}^2$ always intersect. This kind of phenomenon happens here: the dimension of the affine space (1.7) forces a lower bound for the codimension of $\mathcal{P}_m(A)$. More precisely, applying a result from [BG1] we obtain:

$$\operatorname{codim} \mathcal{P}_m(A) \geqslant m + 1 - r(A). \tag{1.8}$$

So we need to show that matrices A with large $r(A)$ are rare. A careful matrix analysis provides an upper bound to $r(A)$ based on the numbers and sizes of the Jordan blocks of A, and on the occasional algebraic relations between the eigenvalues. This bound together with (1.8) implies (1.6) and therefore concludes the proof of Theorem 1.9.

In fact, the results of this analysis are even better, and we conclude that the codimension inequality (1.6) is strict when $k \geqslant 1$. This implies that poor data $(A, B_1, \dots, B_m)$ for which the matrix A is degenerate form a subset of $\mathcal{P}_m^{(\mathbb{C})}$ with strictly bigger codimension. Thus we can show that the poor data that appear generically are well-behaved, which leads to Theorem 1.2.

1.5. **Other remarks.** One can also study uniform regularity of periodic inputs of higher period. Using our results for constant inputs, it is not difficult to derive some (non-sharp) dimension bounds for singular periodic inputs. However, for highly resonant non-periodic inputs, we have no idea on how to obtain reasonable dimension estimates.

As mentioned above, in paper [BG3] we have dimension estimates for general inputs. These estimates are basically obtained by avoiding highly resonant inputs (which have large codimension themselves). Thus the results of [BG3] are independent from those of these paper. The proofs there are less involved from the point of view of matrix computations, but use more sophisticated transversality theorems.

Of course, it would be interesting to consider these kind of problems for other Lie groups of matrices, but we won't pursue this issue here.

1.6. **Organization of the paper.** Section 2 contains some basic results about transitivity of spaces of matrices and its relation with universal regularity. We also obtain the easy parts of Theorems 1.8 and 1.9, namely (semi)algebraicity and the upper codimension inequalities.

In Section 3 we introduce the concept of rigidity, which is related to the quantity $r(A)$ mentioned above. We state the central rigidity estimates (Theorem 3.7), which consist into two parts. The first and easier part is proved in the same Section 3, while the whole Section 4 is devoted to the proof of the second part.

Section 5 starts with some preliminaries in elementary algebraic geometry. Then we use the rigidity estimates to prove Theorem 1.9, following the strategy outlined above (§ 1.4). Theorem 1.8 follows easily. We also obtain a lemma that is needed for the proof of Theorem 1.2.

In Section 6 we collect some basic facts about stratifications and transversality, and then apply them together with the previous results to obtain Theorems 1.1 and 1.2.

The paper also has three short appendices.

Appendix A basically reobtains the major results in the special case $m = 1$, where we actually gain additional information of practical value: as mentioned in § 1.4, it is possible to describe explicitly what 1-jets the map A should avoid in order to satisfy the conclusions of Theorems 1.1 and 1.2. The arguments necessary for the $m = 1$ case are much simpler and more elementary than those in Sections 3 to 5. Therefore the appendix is also useful to give the reader some intuition about

the general problem, and as a source of examples. Appendix A is written in a slightly informal way, and it can be read after Section 2 (though the final part requires Lemmas 3.1 and 3.2).

In Appendix B we take a closer look to the generic singular constant inputs, and in particular justify Remark 1.3. We also discuss the generic validity of some control-theoretic properties related to accessibility and regularity.

Finally, in Appendix C we apply Theorem 1.9 to prove a version of Theorem 1.1 for holomorphic mappings.

2. Preliminary facts on the poor data

In this section, we review some basic properties related to poorness, and prove the easy inequalities in Theorems 1.8 and 1.9.

2.1. **Transitive spaces.** Let E and F be finite-dimensional vector spaces over the field $\mathbb{K}$. Let $\mathcal{L}(E, F)$ be the space of linear maps from E to F. A vector subspace Λ of $\mathcal{L}(E, F)$ is called *transitive* if for every $v \in E \smallsetminus \{0\}$, we have $\Lambda \cdot v = F$, where $\Lambda \cdot v = \{L(v);\ L \in \Lambda\}$.

Under the identification $\mathcal{L}(\mathbb{K}^m, \mathbb{K}^n) = \mathrm{Mat}_{m\times n}(\mathbb{K})$, we may also speak of transitive spaces of matrices.

Example 2.1. Recall that a *Toeplitz matrix*, resp. a *Hankel matrix*, is a matrix of the form

$$\begin{pmatrix} t_0 & t_1 & \cdots & t_{d-1} \\ t_{-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_1 \\ t_{-d+1} & \cdots & t_{-1} & t_0 \end{pmatrix}, \quad \text{resp.} \quad \begin{pmatrix} h_1 & \cdots & h_{d-1} & h_d \\ \vdots & ⋰ & ⋰ & h_{d+1} \\ h_{d-1} & ⋰ & ⋰ & \vdots \\ h_d & h_{d+1} & \cdots & h_{2d-1} \end{pmatrix},$$

The set of Toeplitz matrices and the set of complex Hankel matrices constitute examples transitive subspaces of $\mathfrak{gl}(d, \mathbb{K})$. Transitivity of the Toeplitz space is a particular case of Example 2.2, and transitivity of Hankel space follows from Remark 2.3. For $\mathbb{K} = \mathbb{C}$, these spaces are optimal, in the sense that they have the least possible dimension; see [Az].

Example 2.2. A *generalized Toeplitz space* is a subspace Λ of $\mathrm{Mat}_{d\times d}(\mathbb{K})$ (where $d \geqslant 2$) with the following property: For any two matrix entries (i_1, j_1) and (i_2, j_2) which are not in the same diagonal (i.e., $i_1 - j_1 \neq i_2 - j_2$), the linear map $(b_{i,j})_{i,j} \in \Lambda \mapsto (b_{i_1,j_1}, b_{i_2,j_2}) \in \mathbb{C}^2$ is onto. Equivalently, a space is generalized Toeplitz if it can be defined by a number of linear relations between the matrix coefficients so that each relation involves only the entries on a same diagonal, and so that the relations do not force any matrix entry to be zero. We will prove later (see § 3.3) that *every generalized Toeplitz space is transitive.*

Remark 2.3. If Λ is a transitive subspace of $\mathcal{L}(E, F)$ and $P \in \mathcal{L}(E, E)$, $Q \in \mathcal{L}(F, F)$ are invertible operators then $P \cdot \Lambda \cdot Q := \{PLQ;\ L \in \Lambda\}$ is a transitive subspace of $\mathcal{L}(E, F)$.

Let us see that transitivity is a semialgebraic or algebraic property, according to the field. Recall that:

- A subset of $\mathbb{K}^n$ is called *algebraic* if it is expressed by polynomial equations with coefficients in $\mathbb{K}$.
- A subset of $\mathbb{R}^n$ is called *semialgebraic* if it is expressed by finitely many polynomial equations or inequalities with coefficients in $\mathbb{R}$.

Proposition 2.4. *Let* $\mathcal{N}^{(\mathbb{K})}_{m,n,k}$ *be the set of* $(B_1, \dots, B_k) \in [\mathrm{Mat}_{m\times n}(\mathbb{K})]^k = \mathbb{K}^{mnk}$ *such that* $\mathrm{span}\{B_1, \dots, B_k\}$ *is not transitive. Then:*

1. *The set* $\mathcal{N}^{(\mathbb{R})}_{m,n,k}$ *is semialgebraic.*
2. *The set* $\mathcal{N}^{(\mathbb{C})}_{m,n,k}$ *is algebraic.*

Proof. [9] Consider the set of $(B_1, \dots, B_k, v) \in [\mathrm{Mat}_{m\times n}(\mathbb{K})]^k \times \mathbb{K}^n_*$ such that

$$\mathrm{span}\{B_1, \dots, B_k\} \cdot v \neq \mathbb{K}^m .$$

This is an algebraic set, because it is expressed by the vanishing of certain determinants. Taking $\mathbb{K} = \mathbb{R}$ and projecting this set along the $\mathbb{R}^n_*$ fiber we obtain $\mathcal{N}^{(\mathbb{R})}_{m,n,k}$; so, by the Tarski–Seidenberg theorem (see [BCR, p. 26]), this set is semialgebraic, proving part 1.

To see part 2, we take $\mathbb{K} = \mathbb{C}$ and projectivize the $\mathbb{C}^n_*$ fiber, obtaining an algebraic subset $[\mathrm{Mat}_{m\times n}(\mathbb{C})]^k \times \mathbb{C}\mathrm{P}^{n-1}$ whose projection along the $\mathbb{C}\mathrm{P}^{n-1}$ fiber is $\mathcal{N}^{(\mathbb{C})}_{m,n,k}$. So part 2 follows from the fact that projections along projective fibers are closed maps with respect to the Zariski topology (see Proposition 5.1 below). ☐

Another important fact is that complex transitivity of real matrices is a stronger property than real transitivity:

Proposition 2.5. *The real part of* $\mathcal{N}^{(\mathbb{C})}_{m,n,k}$ *(that is, its intersection with* $[\mathrm{Mat}_{m\times n}(\mathbb{R})]^k$*) contains* $\mathcal{N}^{(\mathbb{R})}_{m,n,k}$.

Moreover, the inclusion can be strict. The explanation is this: real matrix data can be $\mathbb{R}$-transitive without being $\mathbb{C}$-transitive because the directions that detect non-transitivity are non-real. A formal proof and examples are provided in [BG2].

Remark 2.6. The codimension of $\mathcal{N}^{(\mathbb{C})}_{m,n,k}$ is computed in [BG2]: it is $\max(k-m-n+2, 0)$. We also observe in [BG2] that $\mathcal{N}^{(\mathbb{R})}_{m,n,k}$ can fail to be real-algebraic. But we will not need those results in the present paper.

2.2. **Universal regularity for constant inputs and richness.** In this subsection we prove Proposition 1.6; in fact we prove a more precise result, and also fix some notation.

Recall that if $A \in \mathrm{GL}(d, \mathbb{K})$ then the *adjoint of* A is the linear operator Ad_A on $\mathfrak{gl}(d, \mathbb{K})$ given by the formula $\mathrm{Ad}_A(B) = ABA^{-1}$.

If $A\colon \mathcal{U} \to \mathrm{GL}(d, \mathbb{C})$ is a differentiable map then the *normalized derivative* of A at a point u is the linear map $T_u\mathcal{U} \to \mathfrak{gl}(d, \mathbb{R})$ given by $h \mapsto (DA(u) \cdot h) \circ A^{-1}(u)$.

Let $\phi_N(\xi_0, \hat{u})$ be the state $\xi_N \in \mathbb{K}\mathrm{P}^d$ of the system (1.4) determined by the initial state ξ_0 and the input sequence $\hat{u} \in \mathcal{U}^N$. Let $\partial_2\phi_N(\xi_0, \hat{u})$ be the derivative of the map $\phi_N(\xi_0, \cdot)$ at $\hat{u}$.

Fix a constant input $\hat{u} = (u, \dots, u) \in \mathcal{U}^N$, and local coordinates on $\mathcal{U}$ around u. Let B_j be the normalized partial derivatives of the map A at u with respect to the i^{th} coordinate. Consider the data $\mathbf{A} = (A, B_1, \dots, B_m)$, where $A = A(u)$. Define the following subspace of $\mathfrak{gl}(d, \mathbb{K})$:

$$\Lambda_N(\mathbf{A}) = \mathbb{K} \cdot \mathrm{Id} + \operatorname*{span}_{\substack{0 \leqslant n < N-1 \\ 1 \leqslant j \leqslant m}} \left\{\mathrm{Ad}^n_A(B_j)\right\}, \tag{2.1}$$

Proposition 2.7. *For all* $\xi_0 \in \mathbb{K}\mathrm{P}^{d-1}$ *and any* $x_0 \in \mathbb{K}^d \smallsetminus \{0\}$ *representing* ξ_0,

$$\mathrm{rank}\, \partial_2\phi_N(\xi_0, \hat{u}) = \dim\left[\Lambda_N(\mathbf{A}) \cdot (A^N x_0)\right] - 1.$$

In particular (since $A = A(u)$ is invertible), the input $\hat{u}$ is universally regular if and only if $\Lambda_N(\mathbf{A})$ is a transitive space, which is the statement of Proposition 1.6.

Proof of Proposition 2.7. Let $\xi_0 = [x_0]$, where $x_0 \in \mathbb{K}^d \smallsetminus \{0\}$. Let $\psi_N(x_0, \hat{u})$ be the final state of the non projectivized system (1.3) determined by the initial state

[9] This proof also appears in [BG2]

x_0 and by the sequence of controls $\hat{u} \in \mathcal{U}^N$. Using local coordinates with u in the origin, we have the following first order approximation for $\hat{u} \simeq 0$:

$$\begin{aligned}\psi_N(x_0, \hat{u}) &\simeq A^N x_0 + \sum_{\substack{1 \leqslant j \leqslant m \\ 0 \leqslant t < N}} u_{t,j} A^{N-t-1} B_j A^{t+1} x_0 \\ &= \left(\mathrm{Id} + \sum_{\substack{1 \leqslant j \leqslant m \\ 0 \leqslant n < N}} u_{N-1-n,j} \mathrm{Ad}_A^n (B_j) \right) x_N ,\end{aligned}$$

where $x_N = \psi_N(x_0, 0) = A^N x_0$. Therefore the image of $\partial_2 \psi_N(x_0, \hat{u})$ is the following subspace of $T_{A^N x_0} \mathbb{K}^d$:

$$V = \left(\operatorname*{span}_{\substack{1 \leqslant j \leqslant m \\ 0 \leqslant n < N}} \mathrm{Ad}_A^n B_j \right) \cdot x_N ,$$

The image of $\partial_2 \phi_N(\xi_0, \hat{u})$ equals $D\pi(x_N)(V)$, where $\pi : \mathbb{K}^d \smallsetminus \{0\} \to \mathbb{K}\mathrm{P}^{d-1}$ is the canonical projection. Notice that $\operatorname{Ker} D\pi(x) = \mathbb{K}x$ for any $x \in \mathbb{K}^d \smallsetminus \{0\}$. It follows that

$$\begin{aligned}\operatorname{rank} \partial_2 \phi_N(\xi_0, \hat{u}) &= \dim \left[D\pi(x_N)(V) \right] \\ &= \dim \big[D\pi(x_N) \big(\mathbb{K}x_N + V \big) \big] = \dim[\mathbb{K}x_N + V] - 1\end{aligned}$$

Since $\mathbb{K}x_N + V = \Lambda_N(\mathbf{A}) \cdot x_N$, the proposition is proved. □

The discussion above motivates the introduction of a more general notation, which will be convenient later. Consider a linear operator $H \colon E \to E$, where E is a finite-dimensional vector space over the field $\mathbb{K}$. Given a vector $v \in E$, the *orbit* of v under H is the set $\{H^n v;\ n \geqslant 0\}$. Denote the space spanned by the orbit by $\operatorname{sorb}_H v$. We have

$$\operatorname{sorb}_H v = \big\{ f(H) \cdot v;\ f \text{ is a polynomial with coefficients in } \mathbb{K} \big\} .$$

It follows from the Cayley–Hamilton theorem that $\operatorname{sorb}_H v$ is the space spanned by the first $\dim E$ iterates of v:

$$\operatorname{sorb}_H v = \operatorname{span}\{H^n v;\ n = 0, \ldots, \dim E - 1\}.$$

Let us also denote

$$\operatorname{sorb}_H(v_1, \ldots, v_n) = \operatorname{sorb}_H v_1 + \cdots + \operatorname{sorb}_H v_n .$$

In this notation, the union $\Lambda(\mathbf{A}) := \bigcup_N \Lambda_N(\mathbf{A})$ of the elements of the sequence (2.1) is expressed as

$$\Lambda(\mathbf{A}) = \operatorname{sorb}_{\mathrm{Ad}_A}(\mathrm{Id}, B_1, \ldots, B_m), \quad \text{where } \mathbf{A} = (A, B_1, \ldots, B_m). \tag{2.2}$$

We have $\Lambda_N(\mathbf{A}) = \Lambda(\mathbf{A})$ for all $N \geqslant d^2$, as stated in Remark 1.7.

2.3. **The sets of poor data.** For emphasis, we repeat the definition already gave at the introduction: The data $\mathbf{A} = (A, B_1, \ldots, B_m) \in \mathrm{GL}(d, \mathbb{K}) \times [\mathfrak{gl}(d, \mathbb{K})]^m$ is *rich* if the space $\Lambda(\mathbf{A})$ defined by (2.2) is transitive, and *poor* otherwise. The concept in fact depends on the field under consideration. The set of such poor data is denoted by $\mathcal{P}_{m,d}^{(\mathbb{K})}$.

It follows immediately from Proposition 2.4 that $\mathcal{P}_{m,d}^{(\mathbb{R})}$ is a closed and semialgebraic subset of $\mathrm{GL}(d, \mathbb{R}) \times [\mathfrak{gl}(d, \mathbb{R})]^m$ and $\mathcal{P}_{m,d}^{(\mathbb{C})}$ is an algebraic subset of $\mathrm{GL}(d, \mathbb{C}) \times [\mathfrak{gl}(d, \mathbb{C})]^m$. This proves part of Theorems 1.8 and 1.9.

Also, by Proposition 2.5 the real poor data are contained in the real part of the complex poor data, i.e.,

$$\mathcal{P}_{m,d}^{(\mathbb{R})} \cap \big[\mathrm{GL}(d, \mathbb{K}) \times [\mathfrak{gl}(d, \mathbb{K})]^m \big] \subset \mathcal{P}_{m,d}^{(\mathbb{C})} . \tag{2.3}$$

Let us also note that the sets of poor data are saturated in the sense of the following definition: A set $\mathcal{Z} \subset [\mathrm{Mat}_{d\times d}(\mathbb{K})]^{1+m}$ will be called *saturated* if $(A, B_1, \dots, B_m) \in \mathcal{Z}$ implies that:

- For all $P \in \mathrm{GL}(d, \mathbb{K})$, the tuple $(P^{-1}AP, P^{-1}B_1P, \dots, P^{-1}B_mP)$ belongs to $\mathcal{Z}$.
- For all $Q = (q_{ij}) \in \mathrm{GL}(m, \mathbb{K})$, the tuple $(A, B_1', \dots, B_m')$, where $B_i' = \sum_j q_{ij}B_j$, belongs to $\mathcal{Z}$.

Remark 2.8. 1. A subset $[\mathrm{Mat}_{d\times d}(\mathbb{K})]^{1+m}$ is saturated if and only if it is invariant under a certain action of the group $\mathrm{GL}(d, \mathbb{K}) \times \mathrm{GL}(m, \mathbb{K})$.
2. The real part of a complex saturated set is saturated (in the real sense).

2.4. **The easy codimension inequality of Theorems 1.8 and 1.9.** Here we will discuss the simplest examples of poor data.

To begin, notice that if $A \in \mathrm{GL}(d, \mathbb{C})$ is diagonalizable then so is Ad_A. Indeed, assume without loss of generality that $A = \mathrm{Diag}(\lambda_1, \dots, \lambda_d)$. Consider the basis $\{E_{i,j};\ i, j \in \{1, \dots, d\}\}$ of $\mathfrak{gl}(d, \mathbb{C})$, where

$$E_{i,j} \text{ is the matrix whose only nonzero entry is a 1 in the } (i, j) \text{ position.} \tag{2.4}$$

Then $\mathrm{Ad}_A(E_{i,j}) = \lambda_i\lambda_j^{-1}E_{i,j}$. We summarize this fact as:

$$\mathrm{Ad}_A = \mathrm{Diag}\begin{pmatrix} 1 & \lambda_1\lambda_2^{-1} & \cdots \\ \lambda_2\lambda_1^{-1} & 1 & \\ \vdots & & \ddots \end{pmatrix}. \tag{2.5}$$

So if f is a polynomial and $B = (b_{ij})$ then

$$\text{the } (i, j)\text{-entry of the matrix } (f(\mathrm{Ad}_A))(B) \text{ is } f(\lambda_i\lambda_j^{-1})b_{ij}. \tag{2.6}$$

The data $\mathbf{A} = (A, B_1, \dots, B_m) \in \mathrm{GL}(d, \mathbb{K}) \times \mathfrak{gl}(d, \mathbb{K})^m$ is called *conspicuously poor* if there exists a change of bases $P \in \mathrm{GL}(d, \mathbb{K})$ such that:

- the matrix $P^{-1}AP$ is diagonal;
- the matrices $P^{-1}B_kP$ have a zero entry in a common off-diagonal position; more precisely, there are indices $i_0, j_0 \in \{1, \dots, d\}$ with $i_0 \neq j_0$ such that for each $k \in \{1, \dots, m\}$, the (i_0, j_0) entry of the matrix $P^{-1}B_kP$ vanishes.

(As in the definition of poorness, the concept depends on the field $\mathbb{K}$.)

Lemma 2.9. *Conspicuously poor data are poor.*

Proof. Let $\mathbf{A} = (A, B_1, \dots, B_m)$ be conspicuously poor. With a change of basis we can assume that A is diagonal. Let $(e_1, \dots, e_d)$ be the canonical basis of $\mathbb{K}^d$. Let (i, j) be the entry position where all B_i's have a zero entry. By (2.6), all matrices in the space $\Lambda(\mathbf{A})$ given by (2.2) have a zero entry in the (i_0, j_0) position. In particular, there is no $L \in \Lambda(\mathbf{A})$ such that $L \cdot e_{j_0} = e_{i_0}$, showing that this space is not transitive. □

The converse of this lemma is certainly false. (Many examples appear in Appendix A; see also Example 3.6.) However, we will see in § 2.5 that the converse holds for generic A.

We will use Lemma 2.9 to prove the easy codimension inequalities for Theorems 1.8 and 1.9; first we need to recall the following[10]:

[10] Proposition 2.10 follows from the implicit function theorem; for a proof using complex analysis, see [Ka, p. 67].

Proposition 2.10. *Suppose $A \in \mathrm{Mat}_{d\times d}(\mathbb{K})$ is diagonalizable over $\mathbb{K}$ and with simple eigenvalues only. Then there is a neighborhood of A where the eigenvalues vary smoothly, and where the eigenvectors can be chosen to vary smoothly.*

Proposition 2.11 (Easy half of Theorems 1.8 and 1.9)**.** *For both $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, we have $\operatorname{codim}_{\mathbb{K}} \mathcal{P}_m^{(\mathbb{K})} \leqslant m$.*

Proof. Using Proposition 2.10, we can exhibit smoothly embedded disks of codimension m inside $\mathrm{GL}(d,\mathbb{K}) \times \mathfrak{gl}(d,\mathbb{K})^m$ formed by conspicuously poor data. ☐

2.5. **Unconstrained matrices.** The material from this subsection is used in the proof of Theorem 1.2, but not in the proof of Theorem 1.1. It is also used in Appendix A.

If p is an irreducible factor of the polynomial $\lambda_i\lambda_\ell - \lambda_j\lambda_k$ then the relation $p = 0$ is called an *elementary constraint* in the variables $\lambda_1, \dots, \lambda_d$. Every elementary constraint can be written, after a permutation of the indices $1, \dots, d$, as one of the following:

- a type 1 constraint: $\lambda_1\lambda_3 = \lambda_2^2$.
- a type 2 constraint: $\lambda_1\lambda_4 = \lambda_2\lambda_3$.
- a type 3 constraint: $\lambda_1 = -\lambda_2$.
- a type 4 constraint: $\lambda_1 = \lambda_2$.

We say that a matrix $A \in \mathrm{GL}(d,\mathbb{R})$ is *unconstrained* if its eigenvalues, counted with multiplicity, do not satisfy any elementary constraint.

Remark 2.12. A matrix A is unconstrained if and only if Ad_A has the maximal possible number of distinct eigenvalues, namely, $d^2 - d + 1$. This is obvious from (2.5) if one restricts to diagonalizable matrices A. The general case follows from the fact (which we will prove rigorously in § 4.3) that the multiplicities of the eigenvalues of Ad_A are those "predicted" by formula (2.5).

Let us see that the converse of Lemma 2.9 holds for unconstrained A:

Lemma 2.13. *Suppose that the data $\mathbf{A} = (A, B_1, \dots, B_m) \in \mathrm{GL}(d,\mathbb{K}) \times \mathfrak{gl}(d,\mathbb{K})^m$ is poor and that the matrix A is unconstrained. Then $\mathbf{A}$ is conspicuously poor.*

Proof. Suppose A is unconstrained. In particular, A has simple spectrum. With a change of basis we can assume that A is diagonal.

Now suppose that $\mathbf{A} = (A, B_1, \dots, B_m)$ is not conspicuously poor. This means that for each off-diagonal position there is at least of of the matrices B_k that has a non-zero entry in that position. (Notice that this fact does not depend on the change of basis chosen before.)

Since A is unconstrained, the values $\lambda_i\lambda_j^{-1}$, where (i,j) runs on the matrix positions outside the diagonal, are pairwise different, and all different from 1. Recall that one can always (using Lagrange formula) find a polynomial whose values at finitely many different points are prescribed. So It follows from (2.6) that the space $\Lambda(\mathbf{A})$ contains all matrices (y_{ij}) such that $y_{11} = \cdots = y_{dd}$, and in particular, all Toeplitz matrices. So $\Lambda(\mathbf{A})$ is transitive, i.e., $\mathbf{A}$ is not poor. This proves the lemma. ☐

Let us establish another simple result, which is related to Theorem 1.2. Denote by $(e_1, \dots, e_d)$ the canonical basis of $\mathbb{C}^d$.

Lemma 2.14. *Suppose that the data $\mathbf{A} = (A, B_1, \dots, B_m) \in \mathrm{GL}(d,\mathbb{C}) \times \mathfrak{gl}(d,\mathbb{C})^m$ has the following properties:*

1. *A is an unconstrained diagonal matrix;*
2. *there are indices i_0, $j_0 \in \{1, \dots, d\}$ with $i_0 \neq j_0$ such that for each $k \in \{1, \dots, m\}$, the (i_0, j_0) entry of the matrix B_k vanishes;*

3. *the off-diagonal vanishing entry position* (i_0, j_0) *above is unique.*

Then:

1. *There is a single direction* $[v] \in \mathbb{C}\mathrm{P}^{d-1}$ *such that* $\Lambda(\mathbf{A})\cdot v \neq \mathbb{C}^d$*, namely* $[e_{j_0}]$.
2. *The space* $\Lambda(\mathbf{A}) \cdot e_{j_0}$ *has codimension* 1*; in fact, it equals* $\operatorname{span}\{e_i;\ i \neq i_0\}$.

Proof. Under the assumptions on $\mathbf{A}$, the space $\Lambda(\mathbf{A})$ contains

$$\big\{(y_{ij}) \in \mathfrak{gl}(d,\mathbb{C});\ y_{11} = \cdots y_{dd},\ y_{i_0 j_0} = 0\big\}.$$

The conclusions follow easily. □

After the preliminaries above, the optional Appendix A can be read (as we mentioned in § 1.6).

3. Rigidity

The aim of this section is to state Theorem 3.7 and prove its first part. Along the way we will establish several lemmas which will be reused in the proof of the second part of the theorem in Section 4.

3.1. **Acyclicity.** Consider a linear operator $H\colon E \to E$, where E is a finite-dimensional complex vector space.

The operator H is called *cyclic* if it has a *cyclic vector*, that is, some $v \in E$ such that $\operatorname{sorb}_H v$ is the whole space E. The following two lemmas are useful to find cyclic vectors, when they exist:

Lemma 3.1. *Suppose that* $E = \mathbb{C}^\ell$ *and that* H *is a Jordan block:*

$$H = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}.$$

Then a vector $v = (x_1, \dots, x_\ell)$ *is cyclic for* H *if and only if* $x_\ell \neq 0$.

Proof. For any polynomial f we have (see [Ga, page 100]):

$$f(H) = \begin{pmatrix} f(\lambda) & \frac{f'(\lambda)}{1!} & \frac{f''(\lambda)}{2!} & \cdots & \frac{f^{(\ell-1)}(\lambda)}{(\ell-1)!} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \frac{f''(\lambda)}{2!} \\ & & & \ddots & \frac{f'(\lambda)}{1!} \\ & & & & f(\lambda) \end{pmatrix}.$$

So the space spanned by the powers of H is the space of upper triangular Toeplitz matrices. The rest of the proof is an easy exercise. □

Lemma 3.2. *Let* E *be a finite-dimensional complex vector space and let* $H\colon E \to E$ *be a linear operator. Assume that* $E_1, \dots, E_k \subset E$ *are* H*-invariant subspaces and that the spectra of* $A|E_i$ $(1 \leqslant i \leqslant k)$ *are pairwise disjoint. If* $v_1 \in E_1, \dots, v_k \in E_k$ *then*

$$\operatorname{sorb}_H(v_1, \dots, v_k) = \operatorname{sorb}_H(v_1 + \cdots + v_k)\,.$$

Proof. The $\supset$ part is trivial; let us show the $\subset$ part. Take $w \in \operatorname{sorb}_H(v_1, \dots, v_k)$, so $w = \sum f_i(H) \cdot v_i$, where each f_i is a polynomial. Let p_i be the minimal polynomial of $H|E_i$, and let $q_i = \prod_{j \neq i} p_j$. Since the spectra of $A|E_i$ are pairwise disjoint, the polynomials p_i are pairwise relatively prime, and so the polynomials q_i are jointly relatively prime. Since polynomials form a principal ideal domain, there exist polynomials g_i such that $\sum g_i q_i = 1$. Using that $q_i(H) \cdot v_j = 0$ if $i \neq j$, we have:

$$w = \sum_i f_i(H) \cdot v_i = \sum_i f_i(H) \left(\sum_j g_j(H) q_j(H) \right) \cdot v_i$$
$$= \sum_i f_i(H) g_i(H) q_i(H) \cdot v_i = \left(\sum_i f_i(H) g_i(H) q_i(H) \right) \cdot \sum_j v_j .$$

That is, $w = f(H) \cdot \sum_j v_j$ for some polynomial f, as we wanted to show. $\square$

We define the *acyclicity* of H as the least number n of vectors $v_1, \dots, v_n \in E$ such that $\operatorname{sorb}_H(v_1, \dots, v_n) = E$. We denote $n = \operatorname{acyc} H$. So $\operatorname{acyc} H = 1$ means that H is a cyclic operator.

Let us relate acyclicity with the Jordan normal form of H. The *geometric multiplicity* of an eigenvalue λ of H is the number of corresponding Jordan blocks or, equivalently, the dimension of the kernel of $H - \lambda \mathrm{Id}$. The following fact is probably well-known, but since we could not find a precise reference we provide a proof:[11]

Proposition 3.3. *The acyclicity of an operator equals the maximum of the geometric multiplicities of its eigenvalues.*

Proof. Let $\lambda_1, \dots, \lambda_k$ be the eigenvalues of H, counted without multiplicity, and $E = E_1 \oplus \cdots \oplus E_k$ be the splitting into generalized eigenspaces. Let n_i be the geometric multiplicity of λ_i, and let $n = \max n_i$.

Using Lemma 3.1, we find $v_{i,1}, \dots, v_{i,n_i} \in E_i$ such that $\operatorname{sorb}_H(v_{i,1}, \dots, v_{i,n_i}) = E_i$. Define $v_{i,j} = 0$ for $n_i < j \leqslant n$. Consider $w_j = \sum_{i=1}^k v_{i,j}$, for $j = 1, \dots, n$. By Lemma 3.2, $\operatorname{sorb}_H w_j = \operatorname{sorb}_H(v_{1,j}, \dots, v_{k,j})$. So

$$\operatorname{sorb}_H(w_1, \dots, w_n) = \sum_j \operatorname{sorb}_H(v_{1,j}, \dots, v_{k,j}) = \sum_i \operatorname{sorb}_H(v_{i,1}, \dots, v_{i,n}) = E.$$

This shows that $\operatorname{acyc} H \leqslant n$.

To show the reverse inequality, assume that $n = n_1$, for example. For each vector in E, write its coordinates with respect to the Jordan basis, and the consider only the coordinates corresponding to the rightmost columns of the Jordan blocks for λ_1. This defines a linear map $P \colon E \to \mathbb{C}^n$ such that $PH = \lambda_1 P$. Now take any vectors $u_1, \dots, u_{n-1} \in E$. Then the space $S = \operatorname{sorb}_H(u_1, \dots, u_{n-1})$ is sent by P to the vector space $\operatorname{span}\{Pu_1, \dots, Pu_{n-1}\}$, which has dimension $\leqslant n$. Since P is onto $\mathbb{C}^n$, the space S cannot be the whole E. This shows that $\operatorname{acyc} H \geqslant n$, completing the proof. $\square$

Remark 3.4. The operators which interest us most are $H = \mathrm{Ad}_A$, where $A \in \mathrm{GL}(d, \mathbb{C})$. It is useful to observe that *the geometric multiplicity of* 1 *as an eigenvalue of* Ad_A *equals the the codimension of the conjugacy class of* A *inside* $\mathrm{GL}(d, \mathbb{C})$. To prove this, consider the map $\Psi_A \colon \mathrm{GL}(d, \mathbb{C}) \to \mathrm{GL}(d, \mathbb{C})$ given by $\Psi_A(X) = \mathrm{Ad}_X(A)$. The derivative at $X = \mathrm{Id}$ is $H \mapsto HA - AH$; so $\operatorname{Ker} D\Psi_A(\mathrm{Id}) = \operatorname{Ker}(\mathrm{Ad}_A - \mathrm{id})$. Therefore when $X = \mathrm{Id}$, the rank of $D\Psi_A(X)$ equals the geometric multiplicity of 1 as an eigenvalue of Ad_A. To see that this is

[11]The usual textbook approach is the other way around: one uses results about cyclic operators to obtain the Jordan normal form; see e.g. [Ga].

true for any X, notice that $\Psi_A = \Psi_{\mathrm{Ad}_X(A)} \circ R_{X^{-1}}$ (where R denotes a right-multiplication diffeomorphism of $\mathrm{GL}(d,\mathbb{C})$).

We will see later (Lemma 4.15) that 1 is the eigenvalue of Ad_A with the biggest geometric multiplicity. By Proposition 3.3, we conclude that $\operatorname{acyc}\mathrm{Ad}_A$ equals the codimension of the conjugacy class of A.

3.2. **Definition of rigidity, and the main rigidity estimate.** Let E and F be finite-dimensional complex vector spaces. Let H be a linear operator action on the space $\mathcal{L}(E,F)$. We define the *rigidity* of H, denoted $\operatorname{rig} H$, as the least n such that there exist $L_1, \dots, L_n \in \mathcal{L}(E,F)$ so that $\operatorname{sorb}_H(L_1,\dots,L_n)$ is transitive. Therefore

$$1 \leqslant \operatorname{rig} H \leqslant \operatorname{acyc} H\,.$$

For technical reasons, we also define a *modified rigidity* of H, denoted $\operatorname{rig}_+ H$. The definition is the same, with the difference that if $E = F$ then L_1 is required to be the identity map in $\mathcal{L}(E,E)$. Of course,

$$\operatorname{rig} H \leqslant \operatorname{rig}_+ H \leqslant \operatorname{rig} H + 1.$$

We want to give a reasonably good estimate of the modified rigidity of Ad_A for any fixed $A \in \mathrm{GL}(d,\mathbb{C})$. (This will be achieved in Lemma 4.18.) We assume that $d \geqslant 2$; so $\operatorname{rig}_+ \mathrm{Ad}_A \geqslant 2$. The next example shows that "most" matrices A have the lowest possible $\operatorname{rig}_+ \mathrm{Ad}_A$.

Example 3.5. If $A \in \mathrm{GL}(d,\mathbb{C})$ is unconstrained (see § 2.5) then $\operatorname{rig}_+ \mathrm{Ad}_A = 2$. Indeed if we take a matrix $B \in \mathfrak{gl}(d,\mathbb{C})$ whose expression in the base that diagonalizes A has no zeros off the diagonal then, by Lemma 2.13, $\Lambda(A,B) = \operatorname{sorb}_{\mathrm{Ad}_A}(\mathrm{Id}, B)$ is rich.

More generally, if $A \in \mathrm{GL}(d,\mathbb{C})$ is little constrained (see Appendix A) then it follows from Proposition A.2 that $\operatorname{rig}_+ \mathrm{Ad}_A = 2$.

Example 3.6. Consider $A = \mathrm{Diag}(1,\alpha,\alpha^2)$ where $\alpha = e^{2\pi i/3}$. (In the terminology of § 2.5, A has constraints of type 1.) Since Ad_A^3 is the identity, we have $\dim \operatorname{sorb}_{\mathrm{Ad}_A}(\mathrm{Id}, B) \leqslant 4$ for any $B \in \mathfrak{gl}(3,\mathbb{C})$. By the result of Azoff [Az] already mentioned at Example 2.1, the minimum dimension of a transitive subspace of $\mathfrak{gl}(3,\mathbb{C})$ is 5. This shows that $\operatorname{rig}_+ \mathrm{Ad}_A \geqslant 3$. (Actually, equality holds, as we will see in Example 3.10 below.)

Let T be the set of roots of unity. Define an equivalence relation $\asymp$ on the set $\mathbb{C}^*$ of nonzero complex numbers by:

$$\lambda \asymp \lambda' \iff \lambda/\lambda' \in T. \tag{3.1}$$

We also say that λ, λ' are *equivalent mod* T.

For $A \in \mathrm{GL}(d,\mathbb{C})$, we denote

$$c(A) := \text{number of different classes mod } T \text{ of the eigenvalues of } A. \tag{3.2}$$

We now state a technical result which has a central role in our proofs, as explained informally in § 1.4:

Theorem 3.7. *Let $d \geqslant 2$ and $A \in \mathrm{GL}(d,\mathbb{C})$. Then:*

1. *If $c(A) = d$ then* $\operatorname{rig}_+ \mathrm{Ad}_A = 2$.
2. *If $c(A) < d$ then* $\operatorname{rig}_+ \mathrm{Ad}_A \leqslant \operatorname{acyc}\mathrm{Ad}_A - c(A) + 1$.

Remark 3.8. When $c(A) = d$, we have $\operatorname{acyc}\mathrm{Ad}_A = d$ (this will follow from Lemma 4.15); so the conclusion of part 2 does not hold in this case.

Remark 3.9. The conditions of A being unconstrained and A having $c(A) = d$ both mean that A in "non-degenerate". Both of them imply small rigidity, according to Example 3.5 and part 1 of Theorem 3.7. It is important, however, not to confuse the two properties; in fact, none implies the other.

Example 3.10. Consider again A as in Example 3.6. The eigenvalues of Ad_A are 1, α, and α^2, each with multiplicity 3; so Proposition 3.3 gives $\operatorname{acyc} \mathrm{Ad}_A = 3$. So Theorem 3.7 tell us that $\operatorname{rig}_+ \mathrm{Ad}_A \leqslant 3$, which is actually sharp.

The proof of part 1 of Theorem 3.7 will be given in § 3.5 after a few preliminaries (§§ 3.3 and 3.4). These preliminaries are also used in the proof of the harder part 2, which will be given in Section 4.

3.3. **A criterion for transitivity.** We will show the transitivity of certain spaces of matrices that remotely resemble Toeplitz matrices.

Let t, s be positive integers. Let $\mathcal{R}_1$ be a partition of the interval $[1, t] = \{1, \dots, t\}$ into intervals, and let $\mathcal{R}_2$ be a partition of $[1, s]$ into intervals. Let $\mathcal{R}$ be the product partition. We will be interested in matrices of the following special form:

$$M = (m_{i,j})_{\substack{1 \leqslant i \leqslant t \\ 1 \leqslant j \leqslant s}} = \left(\begin{array}{c|c|c} * & 0 & 0 \\ \hline 0 & M_{\mathtt{R}} & 0 \\ \hline 0 & 0 & * \end{array}\right), \tag{3.3}$$

where $\mathtt{R}$ is an element of the product partition $\mathcal{R}$, and $M_{\mathtt{R}}$ is the submatrix $(m_{i,j})_{(i,j) \in \mathtt{R}}$.

Let Λ be a vector space of $t \times s$ matrices. For each $\mathtt{R} \in \mathcal{R}$, say of size $k \times \ell$, we define the following space of matrices:

$$\Lambda^{[\mathtt{R}]} = \big\{ N \in \mathrm{Mat}_{k \times \ell}(\mathbb{C}) \, ; \exists \; M \in \Lambda \text{ of the form (3.3) with } M_{\mathtt{R}} = N \big\}. \tag{3.4}$$

We regard Λ as a subspace of $\mathcal{L}(\mathbb{C}^t, \mathbb{C}^s)$. If the rectangle $\mathtt{R}$ is $[p, p+k] \times [q, q+\ell]$, we regard the space $\Lambda^{[\mathtt{R}]}$ as a subspace of

$$\mathcal{L}\big(\{0\}^{p-1} \times \mathbb{C}^k \times \{0\}^{t-p-k}, \{0\}^{q-1} \times \mathbb{C}^\ell \times \{0\}^{s-q-\ell}\big).$$

Lemma 3.11. *Assume that $\Lambda^{[\mathtt{R}]}$ is transitive for each $\mathtt{R} \in \mathcal{R}$. Then Λ is transitive.*

An interesting feature of the lemma which will be useful later is that it can be applied recursively. Before giving the proof of the lemma, we illustrate its usefulness by showing the transitivity of generalized Toeplitz spaces:

Proof of Example 2.2. Consider the partition of $[1, d]^2$ into 1×1 "rectangles". If Λ is a generalized Toeplitz space then $\Lambda^{[\mathtt{R}]} = \mathrm{Mat}_{1\times 1}(\mathbb{C}) = \mathbb{C}$ for each rectangle $\mathtt{R}$. These are transitive spaces, so Lemma 3.11 implies that Λ is transitive. ☐

Before proving Lemma 3.11, notice the following dual characterization of transitivity, whose proof is immediate:

Lemma 3.12. *A subspace $\Lambda \subset \mathcal{L}(\mathbb{C}^t, \mathbb{C}^s)$ is transitive iff for any non-zero vector $u \in \mathbb{C}^t$ and any non-zero linear functional $\phi \in (\mathbb{C}^s)^*$ there exists $M \in \Lambda$ such that $\phi(M \cdot u) \neq 0$.*

Proof of Lemma 3.11. Take any non-zero vector $u = (u_1, \dots, u_t)$ in $\mathbb{C}^t$ and a non-zero functional $\phi(v_1, \dots, v_s) = \sum_{j=1}^s \phi_j v_j$ in $(\mathbb{C}^s)^*$. By Lemma 3.12, we need to show that there exists $M = (x_{ij}) \in \Lambda$ such that

$$\phi(M \cdot u) = \sum_{i=1}^t \sum_{j=1}^s \phi_j x_{ij} u_i \tag{3.5}$$

is non-zero.

Let i_0 be the least index such that $u_i \neq 0$, and let j_0 be the greatest index such that $\phi_j \neq 0$. Let $\mathtt{R}$ be the element of $\mathcal{R}$ that contains (i_0, j_0). Notice that if M is of the form (3.3) then the (i,j)-entries of M that are above left (resp. below right) of $\mathtt{R}$ do not contribute to the sum (3.5), because ϕ_i (resp. u_j) vanishes. That is, $\phi(M \cdot u)$ depends only on $M_{\mathtt{R}}$ and is given by $\sum_{(i,j)\in\mathtt{R}} \phi_j x_{ij} u_i$; Since $\Lambda^{[\mathtt{R}]}$ is transitive, by Lemma 3.12 there is a choice of a matrix $M \in \Lambda$ of the form (3.3) so that $\phi(M \cdot u) \neq 0$. So we are done. □

3.4. **Preorder in the complex plane.** We consider the set $\mathbb{C}_*/T$ of equivalence classes of the relation (3.1). Since T is the torsion subgroup of $\mathbb{C}_*$, the quotient $\mathbb{C}_*/T$ is an abelian torsion-free group. Therefore it admits a multiplication-invariant total order $\preccurlyeq$, by a result of Levi [Le].[12]

Let $[z] \in \mathbb{C}_*/T$ denote the equivalence class of $z \in C_*$. Let us extend the notation, writing $z \preccurlyeq z'$ if $[z] \preccurlyeq [z']$. Then $\preccurlyeq$ becomes a multiplication-invariant total preorder on $\mathbb{C}_*$ that induces the equivalence relation $\asymp$. In other words, for all z, z', $z'' \in \mathbb{C}_*$ we have:

- $z \preccurlyeq z'$ or $z' \preccurlyeq z$;
- $z \preccurlyeq z'$ and $z' \preccurlyeq z \iff z \asymp z'$;
- $z \preccurlyeq z'$ and $z' \preccurlyeq z'' \implies z \preccurlyeq z''$;
- $z \preccurlyeq z' \implies zz'' \preccurlyeq z'z''$.

It follows that:

- $z \preccurlyeq z' \implies (z')^{-1} \preccurlyeq z^{-1}$.

We write $z \prec z'$ when $z \preccurlyeq z'$ and $z \not\asymp z'$.

3.5. **Proof of the easy part of Theorem 3.7.**

Proof of part 1 of Theorem 3.7. If $c(A) = d$ then in particular all eigenvalues are different and so the matrix A is diagonalizable. So with a change of basis we can assume that $A = \mathrm{Diag}(\lambda_1, \dots, \lambda_d)$. We can also assume that the eigenvalues are increasing with respect to the preorder introduced in § 3.4:

$$\lambda_1 \prec \lambda_2 \prec \cdots \prec \lambda_d \,.$$

Fix any matrix B with only nonzero entries, and consider the space $\Lambda = \mathrm{sorb}_{\mathrm{Ad}_A} X$, which is described by (2.6). We will use Lemma 3.11 to show that Λ is transitive. Let $\mathcal{R}$ be the partition of $[1,d]^2$ into 1×1 rectangles. Given a cell $\mathtt{R} = \{(i_0, j_0)\} \in \mathcal{R}$ and a coefficient $t \in \mathbb{C}$, there exists a polynomial f such that $f(\lambda_i \lambda_j^{-1})$ equals t if $\lambda_i \lambda_j^{-1} = \lambda_{i_0} \lambda_{j_0}^{-1}$ and equals 0 otherwise. Because the eigenvalues are ordered, $M = f(\mathrm{Ad}_A) \cdot B$ is a matrix in Λ of the form (3.3). Also, $M_{\mathtt{R}} = (t)$. So $\Lambda^{[\mathtt{R}]} = \mathbb{C}$, which is transitive. This shows that $\mathrm{rig}\, \mathrm{Ad}_A = 1$, and $\mathrm{rig}_+ \mathrm{Ad}_A \leqslant 2$. Thus, as $d \geqslant 2$, we have $\mathrm{rig}_+ \mathrm{Ad}_A = 2$. □

4. Proof of the hard part of the rigidity estimate

This section is wholly devoted to prove part 2 of Theorem 3.7. In the course of the proof we need to introduce some terminology and to establish several intermediate results. None of these are used in the rest of the paper, apart form a simple consequence, which is Remark 4.16.

---

[12] Let us give a direct proof of the existence of an invariant order on $\mathbb{C}_*/T$. There is an isomorphism between $\mathbb{R} \oplus (\mathbb{R}/\mathbb{Q})$ and $\mathbb{C}_*/T$, namely $(x, y) \mapsto \exp(x + 2\pi i y)$. So it suffices to find an invariant order in $\mathbb{R}/\mathbb{Q}$ (and then take the lexicographic order). Take a Hamel basis B of the $\mathbb{Q}$-vector space $\mathbb{R}$ so that $1 \in B$. Then $\mathbb{R}/\mathbb{Q}$ is a direct sum of abelian groups $\bigoplus_{x \in B,\ x \neq 1} x\mathbb{Q}$. Order each $x\mathbb{Q}$ in the usual way, take any total order on B, and consider the induced lexicographic order on $\mathbb{R}/\mathbb{Q}$.

4.1. **The normal form.** Let $A \in \mathrm{GL}(d, \mathbb{C})$. In order to describe the estimate on $\mathrm{rig}_+ \mathrm{Ad}_A$, we need to put A in a certain normal form, which we now explain.

We fix a preorder $\preccurlyeq$ on $\mathbb{C}_*$ as in § 3.4.

Let $\lambda_1, \ldots, \lambda_r$ be the eigenvalues of A, listed without repetitions, and with respective multiplicities $s_1, \ldots, s_r$. Assume they are ordered:

$$\lambda_1 \preccurlyeq \cdots \preccurlyeq \lambda_r \, . \tag{4.1}$$

Reindex the sequence of eigenvalues $\lambda_1, \ldots, \lambda_r$ as

$$\lambda_{1,1} \asymp \lambda_{1,2} \asymp \cdots \asymp \lambda_{1,r_1} \prec \lambda_{2,1} \asymp \lambda_{2,2} \asymp \cdots \asymp \lambda_{2,r_2} \prec \cdots$$

Write each eigenvalue in polar coordinates:

$$\lambda_{i,j} = r_i \exp(\theta_{i,j}\sqrt{-1}), \quad \text{where } r_i > 0 \text{ and } 0 \leqslant \theta_{i,j} < 2\pi.$$

Reorder the eigenvalues so that, for each i,

$$\theta_{i,1} < \theta_{i,2} < \cdots < \theta_{i,r_i} \, .$$

With a change of basis, we can assume that A has *modified Jordan form*:

$$A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix}, \quad A_k = \begin{pmatrix} \lambda_k D_{t_{k,1}} & & \\ & \ddots & \\ & & \lambda_k D_{t_{k,\tau_k}} \end{pmatrix}, \tag{4.2}$$

where $t_{k,1} + \cdots + t_{k,\tau_k} = s_k$ and D_t is the following $t \times t$ Jordan block:

$$D_t = \begin{pmatrix} 1 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 1 \end{pmatrix}. \tag{4.3}$$

The matrix A will be fixed from now on.

4.2. **Geography.** This subsection contains several definitions which will be fundamental in all arguments until the end of the section. We will define certain subregions of the set $\{1, \ldots, d\}^2$ of matrix entry positions, which depend on the normal form of the matrix A. Later we will see they are related to Ad_A-invariant subspaces. We will use "geographical" terms for those regions: islands, cities, and districts. The regions will have some numerical attributes (banner, area, population); these attributes may seem mysterious initially, but later we will relate them with numerical invariants of Ad_A (eigenvalues, multiplicities, geometric multiplicities). We also introduce other attributes of the regions (northern and southern cities, latitude of a district) which will be useful later in the proofs of our rigidity estimates.

Recall A is a matrix in normal form as explained in § 4.1. Define three partitions $\mathcal{P}_\mathrm{i}$, $\mathcal{P}_\mathrm{c}$, $\mathcal{P}_\mathrm{d}$ of the set $[1, d] = \{1, \ldots, d\}$ into intervals:

- The partition $\mathcal{P}_\mathrm{i}$ corresponds to equivalence classes of eigenvalues under the relation $\asymp$: the right endpoints of its atoms are the numbers $s_1 + \cdots + s_k$ where $k = r$ or k is such that $\lambda_k \prec \lambda_{k+1}$.
- The partition $\mathcal{P}_\mathrm{c}$ corresponds to eigenvalues: the right endpoints of its atoms are the numbers $s_1 + \cdots + s_k$, where $1 \leqslant k \leqslant r$. So $\mathcal{P}_\mathrm{c}$ refines $\mathcal{P}_\mathrm{i}$.
- The partition $\mathcal{P}_\mathrm{d}$ corresponds to Jordan blocks: the right endpoints of its atoms are the numbers $s_1 + \cdots + s_{k-1} + t_{k,1} + \cdots + t_{k,\ell}$, where $1 \leqslant k \leqslant r$ and $1 \leqslant \ell \leqslant \tau_k$. So $\mathcal{P}_\mathrm{d}$ refines $\mathcal{P}_\mathrm{c}$.

For $* = \mathrm{i}$, c, d, let $\mathcal{P}^2_*$ be the partition of the square $[1,d]^2$ into rectangles that are products of atoms of $\mathcal{P}_*$. The elements of $\mathcal{P}^2_{\mathrm{i}}$ are called *islands*, the elements of $\mathcal{P}^2_{\mathrm{c}}$ are called *cities*, and elements of $\mathcal{P}^2_{\mathrm{d}}$ are called *districts*. Thus the *world* $\mathtt{W} = [1,d]^2$ is a disjoint union of islands, each of them is a disjoint union of cities, each of them is a disjoint union of districts.

Example 4.1. Suppose $d = 17$, A has $r = 5$ eigenvalues

$$\lambda_1 = \exp\tfrac{1}{2}\pi i, \quad \lambda_2 = \exp\tfrac{7}{6}\pi i, \quad \lambda_3 = \exp\tfrac{11}{6}\pi i, \quad \lambda_4 = 2\exp\tfrac{1}{6}\pi i, \quad \lambda_5 = 2\exp\tfrac{5}{6}\pi i$$

with respective Jordan blocks of sizes 4, 2, 1; 3, 2; 2; 2, 1. Then there are 4 islands, 25 cities, and 64 districts. See Fig. 1.

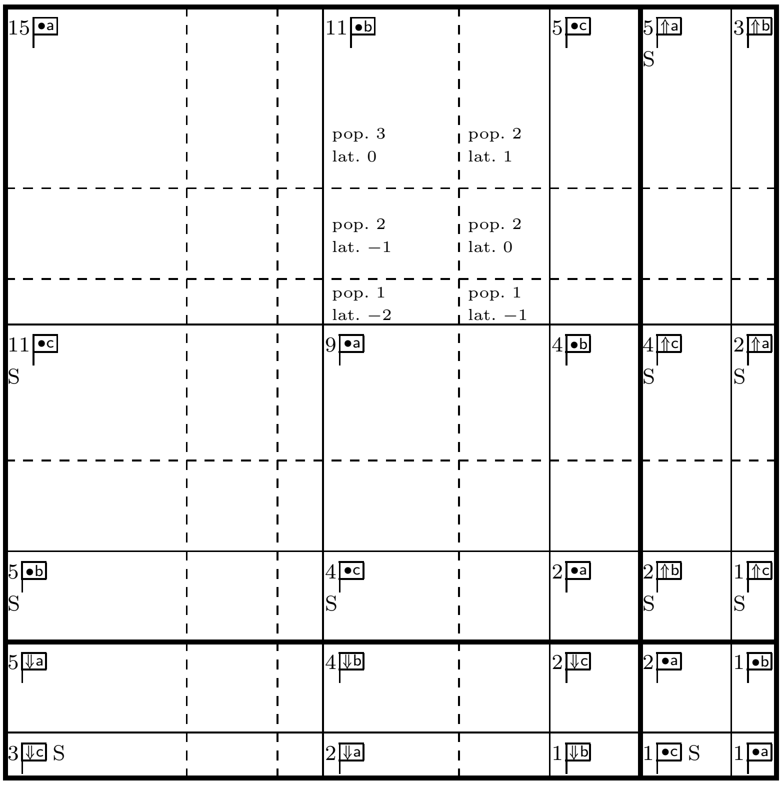

FIGURE 1. The geography corresponding to Example 4.1. Thick (resp., thin, dashed) lines represent island (resp., city, district) borders. Population and latitude of each district inside a selected city are indicated. The population of each city is recorded in its upper left corner, along with a symbolic representation of its banner. There are three banner classes ($\bullet = [1]$, $\Downarrow = [2]$ and $\Uparrow = [1/2]$), each of them with 3 different banners. Southern cities are marked with S.

For each city (or district) we define its *row eigenvalue* and its *column eigenvalue* in the obvious way: If a city $\mathtt{C}$ equals $I_k \times I_\ell$ where I_k and I_ℓ are intervals with right endpoints $s_1 + \cdots + s_k$ and $s_1 + \cdots + s_\ell$, respectively, then the row eigenvalue of $\mathtt{C}$ is λ_k and the column eigenvalue of $\mathtt{C}$ is λ_ℓ. The row and column eigenvalues of a district $\mathtt{D}$ are defined respectively as the row and column eigenvalues of the city that contains $\mathtt{D}$.

Let $\mathtt{C}$ be a city with row eigenvalue $\lambda_{i,j}$ and column eigenvalue $\lambda_{k,\ell}$. The *banner* of $\mathtt{C}$ is defined by $\lambda_{k,\ell}^{-1}\lambda_{i,j}$. The *argument* of the city is the quantity $\theta_{k,\ell} - \theta_{i,j} \in (-2\pi, 2\pi)$. (It coincides, modulo 2π, with the argument of the banner.) The city is

called *southern within its island* if it has strictly negative argument, and *northern within its island* otherwise.

Each district $\mathtt{D}$ has an address of the type "i^{th} row, j^{th} column, city $\mathtt{C}$"; then the *latitude* of the district $\mathtt{D}$ within the city $\mathtt{C}$ is defined as $j-i$. See an example in Fig. 1.

If two cities lie in the same island then their banners are equivalent mod T. Thus every island has a well-defined *banner class* in $\mathbb{C}^*/T$.

If a district, city, or island intersects the diagonal $\{(1,1),\dots,(d,d)\}$ then we call it *equatorial*. Equatorial regions are always square. Thus every equatorial city has banner 1 and every city with banner 1 lies on a equatorial island.

The *area* of a district, city, island or world is defined as the product of its sides. The *population* of a district is defined as the minimum of its sides. Populations of cities, islands and world are defined as the sum of the areas and populations of the corresponding districts.

Let us notice some facts on the location of the banners (which will be useful to apply Lemma 3.11):

Lemma 4.2. *Let* $\mathtt{C}$ *be a city in an island* $\mathtt{I}$*. Consider the divisions of the world* $\mathtt{W}$ *and the island* $\mathtt{I}$ *as in Fig. 2.*

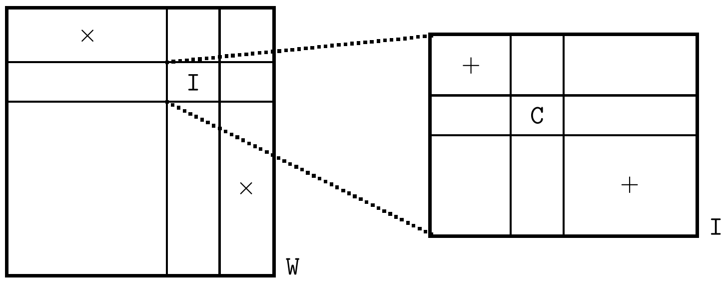

FIGURE 2. The divisions of $\mathtt{W}$ and $\mathtt{I}$ in Lemma 4.2.

Let β be the banner of the city $\mathtt{C}$*, and let $[\beta]$ be the banner class of the island* $\mathtt{I}$*. Then:*

1. *All the islands with banner class $[\beta]$ are inside the regions marked with $\times$.*
2. *If the city* $\mathtt{C}$ *is northern (resp. southern) within* $\mathtt{I}$ *then the all the northern (resp. southern) cities with same banner β are inside the regions marked with $+$.*

Proof. In view of the ordering of the eigenvalues (4.1), the banner class increases strictly (with respect to the order $\prec$, of course) when we move rightwards or upwards to another island. So Claim (1) follows.

The argument of a city takes values in the interval $(-2\pi, 2\pi)$. It increases strictly by moving rightwards or upwards inside $\mathtt{I}$. If two cities in the same island are both northern or both southern then they have the same banner if and only if they have the same argument. So Claim (2) follows. $\square$

4.3. **The adjoint in geographical terms.** Given any $d\times d$ matrix $X=(x_{i,j})$ and a district, city or island $\mathtt{R}=[p,p+t]\times[q,q+s]$ we define the *submatrix* of X corresponding to $\mathtt{R}$ as $(x_{i,j})_{(i,j)\in\mathtt{R}}$. We regard the space of $\mathtt{R}$-submatrices as $\mathcal{L}\big(\{0\}^{p-1}\times\mathbb{C}^t\times\{0\}^{d-p-t}, \{0\}^{q-1}\times\mathbb{C}^s\times\{0\}^{d-q-s}\big)$, or as the set of $d\times d$ matrices whose entries outside $\mathtt{R}$ are all zero. Such spaces are denoted by $\mathtt{R}^\square$, and are invariant under Ad_A. Indeed, if $\mathtt{R}=\mathtt{D}$ is a district then identifying $\mathtt{D}^\square$ with $\mathrm{Mat}_{t\times s}(\mathbb{C})$, the

action of $\mathrm{Ad}_A|\mathtt{D}^\square$ is given by

$$X \mapsto \lambda_k \lambda_\ell^{-1} D_t X D_s^{-1},$$

where $\lambda_k \lambda_\ell^{-1}$ is the banner of $\mathtt{D}$, and D_t, D_s are Jordan blocks defined by (4.3).

If $\mathtt{R}$ is an equatorial district, city, or island we will refer to the $d \times d$-matrix in $\mathtt{R}^\square$ whose $\mathtt{R}$-submatrix is the identity as the *identity on* $\mathtt{R}^\square$. The following observation will be useful:

Lemma 4.3. *If* $\mathtt{D}$ *is an equatorial district then the identity on* $\mathtt{D}^\square$ *is a eigenvalue of the operator* $\mathrm{Ad}_A|\mathtt{D}^\square$ *corresponding to a Jordan block of size* 1×1*.*

Proof. Suppose $\mathtt{D}$ has size $t \times t$. Assume that the claim is false. This means that there exists a matrix $X \in \mathrm{Mat}_{t\times t}(\mathbb{C})$ such that $D_t X D_t^{-1} = X + \mathrm{Id}$, which is impossible because X and $X + \mathrm{Id}$ have different spectra. □

We are going to prove the following:

Lemma 4.4. *For each district* $\mathtt{D}$*, the only eigenvalue of* $\mathrm{Ad}_A|\mathtt{D}^\square$ *is the banner of the city that contains* $\mathtt{D}$*. Moreover, the geometric multiplicity of the eigenvalue is the population of the district.*

The following facts are immediate consequences:

- The eigenvalues of Ad_A are the banners of cities.
- The multiplicity of the eigenvalue β for the operator Ad_A is the total area of cities of banner β.
- The geometric multiplicity of the eigenvalue β for Ad_A is the total population of cities of banner β.

Lemma 4.4 is equivalent to the following:

Lemma 4.5. *Let* $U_{t,s}$ *be the linear operator on* $\mathrm{Mat}_{t\times s}(\mathbb{C})$ *given by*

$$U_{t,s}(X) = D_t X D_s^{-1},$$

where D_t*,* D_s *are Jordan blocks defined by* (4.3)*. Then the only eigenvalue of* $U_{t,s}$ *is* 1*, and its geometric multiplicity is* $\min(t, s)$*.*

The rest of this subsection is devoted to prove Lemma 4.5. To begin, notice that:

$$D_s^{-1} = \begin{pmatrix} 1 & -1 & 1 & \cdots & (-1)^{s-1} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & 1 \\ & & & \ddots & -1 \\ & & & & 1 \end{pmatrix}. \tag{4.4}$$

To describe $U_{t,s}$, it suffices to describe its action on the matrices $E_{i,j}$ whose unique nonzero entry is a 1 in the (i, j) position. Using (4.4), we obtain

$$U_{t,s}(E_{i,j}) = \sum_{p=j}^{s} (-1)^{k-j} (E_{i,p} + E_{i-1,p}) \, ,$$

or, visually:

$$U_{t,s}\left(\quad \boxed{\begin{matrix} & & & & \\ 1 & & & & \end{matrix}} \quad \right) = \left(\quad \boxed{\begin{matrix} 1 & -1 & 1 & \cdots & \pm 1 \\ 1 & -1 & 1 & \cdots & \pm 1 \end{matrix}} \quad \right).$$

The picture above suggests a way of "embedding" all the maps $U_{t,s}$ into a single infinite-dimensional model. More precisely, consider the space $\mathcal{M}$ of infinite matrices of the form $X = (x_{k,\ell})_{k,\ell \leqslant 0}$, where k, ℓ are non-positive integers, that have only finitely many non-zero entries. For each pair of positive integers t, s, define a monomorphism $\iota_{t,s} \colon \mathrm{Mat}_{t\times s}(\mathbb{C}) \to \mathcal{M}$ by

$$(b_{i,j})_{i,j} \mapsto (x_{k,\ell})_{k,\ell} \quad \text{where} \quad x_{k,\ell} = \begin{cases} b_{t+k,s+\ell} & \text{if } k > -t \text{ and } \ell > -s, \\ 0 & \text{otherwise} \end{cases}$$

Define a linear operator $U \colon \mathcal{M} \to \mathcal{M}$ by

$$(x_{k,\ell})_{k,\ell} \mapsto (y_{k,\ell})_{k,\ell} \quad \text{where} \quad y_{k,\ell} = \sum_{q=-\infty}^{\ell} (-1)^{\ell-q}(x_{k,q} + x_{k-1,q}) \,,$$

Then the following diagram commutes:

$$\begin{tikzcd} \mathrm{Mat}_{t\times s}(\mathbb{C}) \arrow[r, "\iota_{t,s}"] \arrow[d, "U_{t,s}"'] & \mathcal{M} \arrow[d, "U"] \\ \mathrm{Mat}_{t\times s}(\mathbb{C}) \arrow[r, "\iota_{t,s}"] & \mathcal{M} \end{tikzcd}$$

Let us prove a few facts about the operator U. It is convenient to consider also $N = U - \mathrm{id}$.

If $X = (x_{k,\ell}) \in \mathcal{M}$ and $n = 1, 2, \ldots$, then we define the n^{th} diagonal of X as the n-tuple $(x_{0,-(n-1)}, x_{-1,-(n-2)}, \ldots, x_{-(n-1),0})$. Define the *height* $h(X)$ of X as 0 if $X = 0$, otherwise $h(X)$ is the maximal n such that X has a nonzero n^{th} diagonal. It is clear that

$$h(N(X)) < h(X) \quad \text{if } X \neq 0. \tag{4.5}$$

It follows that the operator N is nilpotent, in the sense that every orbit eventually hits zero.

Lemma 4.6. *Let $X = (x_{k,\ell}) \in \mathcal{M}$ and let $Z = (z_{k,\ell}) = N(X)$. If $h(X) \leqslant n$ then the n^{th} diagonal of X can be determined from its first element and the $(n-1)^{th}$ diagonal of Z by the formula*

$$x_{-p,-(n-1-p)} = x_{0,-(n-1)} + \sum_{q=0}^{p-1} z_{-q,-(n-2-q)} \,, \quad (p = 0, 1, \ldots, n-1).$$

Proof. It suffices to see that, for each $q = 0, 1, \ldots, n-2$,

$$z_{-q,-(n-2-q)} = x_{-(q+1),-(n-2-q)} - x_{-q,-(n-1-q)} \,. \qquad \square$$

For each $t = 1, 2, \ldots$, let $I_t \in \mathcal{M}$ denote the image under $\iota_{t,t}$ of the $t \times t$ identity matrix. A linear combination of matrices of this type is a matrix with constant diagonals and so will be called a *Toeplitz matrix.*

Lemma 4.7. *$U(X) = X$ if and only if X is a Toeplitz matrix.*

Proof. Since the $t \times t$ identity is fixed by the $U_{t,t}$, we conclude that I_t is fixed by U, proving the "if" part.

To see the converse, take $X = (x_{k,\ell})$ in the kernel of N. Let $n = h(X)$ be height of X. By Lemma 4.6, the n^{th} diagonal of X is constant, say $(c, c, \ldots, c)$. Thus $X - cI_n$ has height at most $n-1$, and belongs to the kernel of N. It follows by induction in n that X is a Toeplitz matrix. $\square$

Proof of Lemma 4.5. Since $U - \mathrm{id}$ is nilpotent, so is $U_{t,s} - \mathrm{id}$, which means that the only eigenvalue of $U_{t,s}$ is 1.

The matrices $I_1, I_2, \ldots, I_{\min(t,s)}$ belong to the image of $\iota_{t,s}$; therefore their inverse images are eigenvectors of $U_{t,s}$. The space V spanned by these eigenvectors is exactly

$$\big\{M \in \mathrm{Mat}_{t\times s}(\mathbb{C});\ \iota_{t,s}(M) \text{ is a Toeplitz matrix}\big\}.$$

By Lemma 4.7, V is also the space of the eigenvectors of $U_{t,s}$. This proves that the geometric multiplicity of $U_{t,s}$ is $\min(t,s)$. □

Remark 4.8. It is natural to ask what are the sizes of the Jordan blocks corresponding to the eigenvectors exhibited in the proof of Lemma 4.4. We don't know the answer, except for the last eigenvector $\iota_{t,s}^{-1}(I_{\min(t,s)})$, which corresponds to a 1×1 Jordan block. This fact, which generalizes Lemma 4.3, can be easily shown using Lemma 4.6.

4.4. **Rigidity estimates for districts and cities.**

Lemma 4.9. *For any district* $\mathtt{D}$*, we have* $\mathrm{rig}_+(\mathrm{Ad}_A|\mathtt{D}^\square) \leqslant \mathrm{pop}\,\mathtt{D}$.

Proof. By Lemma 4.4 (and Proposition 3.3), $\mathrm{Ad}_A|\mathtt{D}^\square$ has acyclicity $n = \mathrm{pop}\,\mathtt{D}$, that is, there are matrices $X_1, \ldots, X_n \in \mathtt{D}^\square$ such that $\mathrm{sorb}_{\mathrm{Ad}_A}(X_1,\ldots,X_n)$ is the whole $\mathtt{D}^\square$ (and, in particular, is transitive in $\mathtt{D}^\square$). So $\mathrm{rig}(\mathrm{Ad}_A|\mathtt{D}^\square) \leqslant n$, which proves the lemma for non-equatorial districts.

If $\mathtt{D}$ is an equatorial district then, by Lemma 4.3, $\mathtt{D}^\square$ splits invariantly into two subspaces, one of them spanned by the the identity matrix on $\mathtt{D}^\square$. So we can choose the matrices X_i above so that X_1 is the identity. This shows that $\mathrm{rig}_+(\mathrm{Ad}_A|\mathtt{D}^\square) \leqslant n$. □

In all that follows, we adopt the convention $\max\varnothing = 0$.

Lemma 4.10. *For any city* $\mathtt{C}$*,*

$$\mathrm{rig}_+(\mathrm{Ad}_A|\mathtt{C}^\square) \leqslant \sum_{\ell \text{ latitude}} \ \max_{\substack{\mathtt{D} \text{ is a district of } \mathtt{C}\\ \text{with latitude } \ell}} \mathrm{rig}_+(\mathrm{Ad}_A|\mathtt{D}^\square)\,.$$

Proof. For each district $\mathtt{D}$ in $\mathtt{C}$, let $r(\mathtt{D}) = \mathrm{rig}_+(\mathrm{Ad}_A|\mathtt{D}^\square)$. Take matrices $X_{\mathtt{D},1}, \ldots, X_{\mathtt{D},r(\mathtt{D})}$ such that $\Lambda_\mathtt{D} := \mathrm{sorb}_{\mathrm{Ad}_A}\big(X_{\mathtt{D},1},\ldots,X_{\mathtt{D},r(\mathtt{D})}\big)$ is a transitive subspace of $\mathtt{D}^\square$, and $X_{\mathtt{D},1}$ is the identity matrix in $\mathtt{D}^\square$ if $\mathtt{D}$ is an equatorial district. Define $X_{\mathtt{D},j} = 0$ for $j > r(\mathtt{D})$. For each latitude ℓ, let n_ℓ be the maximum of $r(\mathtt{D})$ over the districts $\mathtt{D}$ of $\mathtt{C}$ with latitude ℓ, and let

$$Y_{\ell,j} = \sum_{\substack{\mathtt{D} \text{ is a district of } \mathtt{C}\\ \text{with latitude } \ell}} X_{\mathtt{D},j}\,, \quad \text{for } 1 \leqslant j \leqslant n_\ell.$$

Notice that if $\mathtt{C}$ is an equatorial city then $Y_{0,1}$ is the identity matrix in $\mathtt{C}^\square$. Consider the space

$$\Delta = \mathrm{sorb}_{\mathrm{Ad}_A}\big\{Y_{\ell,j};\ \ell \text{ is a latitude, } 1 \leqslant j \leqslant n_\ell\big\}.$$

We claim that for every district $\mathtt{D}$ in $\mathtt{C}$ and for every $M \in \Lambda_\mathtt{D}$, we can find some $N \in \Delta$ with the following properties:

- the submatrix $N_\mathtt{D}$ equals M;
- for every district $\mathtt{D}'$ in $\mathtt{D}$ that has a different latitude than $\mathtt{D}$, the submatrix $N_{\mathtt{D}'}$ vanishes.

Indeed, if $M = \sum_{j=1}^{r(\mathtt{D})} f_j(\mathrm{Ad}_A) X_{\mathtt{D},j}$ for certain polynomials f_j, we simply take $N = \sum_{j=1}^{r(\mathtt{D})} f_j(\mathrm{Ad}_A) Y_{\ell,j}$, where ℓ is the latitude of $\mathtt{D}$.

In notation (3.4), the claim we have just proved means that $\Delta^{[\mathtt{D}]} \supset \Lambda_\mathtt{D}$. So we can apply Lemma 3.11 and conclude that Δ is a transitive subspace of $\mathtt{C}^\square$. Therefore $\mathrm{rig}_+(\mathrm{Ad}_A|\mathtt{C}^\square) \leqslant \sum n_\ell$, as we wanted to show. □

Example 4.11. Using Lemmas 4.9 and 4.10, we see that the city $\mathtt{C}$ whose district populations are indicated in Fig. 1 has $\operatorname{rig}_+(\mathrm{Ad}_A|\mathtt{C}^{\square}) \leqslant 5$.

In fact, we will not use Lemmas 4.9 and 4.10 directly, but only the following immediate consequence:

Lemma 4.12. *For every city* $\mathtt{C}$ *we have* $\operatorname{rig}_+(\mathrm{Ad}_A|\mathtt{C}^{\square}) \leqslant \operatorname{pop}\mathtt{C}$. *The inequality is strict if has more than one row of districts and more that one column of districts.*

4.5. **Comparative demographics.** If $\mathtt{R}$ is a district, city or island, we define its *row projection* $\pi_{\mathrm{r}}(\mathtt{R})$ as the unique equatorial district, city or island (respectively) that is in the same row as $\mathtt{R}$. Analogously, we define the *column projection* $\pi_{\mathrm{c}}(\mathtt{R})$.

Lemma 4.13. *For any city* $\mathtt{C}$, *we have*

$$\operatorname{pop}\mathtt{C} \leqslant \frac{\operatorname{pop}\pi_{\mathrm{r}}(\mathtt{C}) + \operatorname{pop}\pi_{\mathrm{c}}(\mathtt{C})}{2} \,.$$

Moreover, equality implies that the number of rows of districts for $\mathtt{C}$ *equals the number of columns of districts.*

This is a clear consequence of the abstract lemma below, taking x_α, $\alpha \in F_0$ (resp. $\alpha \in F_1$) as the sequence of heights (resp. widths) of districts in $\mathtt{C}$, counting repetitions.

Lemma 4.14. *Let F be a nonempty finite set, and let x_α be positive numbers indexed by $\alpha \in F$. Take any partition $F = F_0 \sqcup F_1$. For ϵ, $\delta \in \{0,1\}$, let*

$$\Sigma_{\epsilon\delta} = \sum_{(\alpha,\beta)\in F_\epsilon \times F_\delta} \min(x_\alpha, x_\beta)\,.$$

Then

$$\Sigma_{01} = \Sigma_{10} \leqslant \frac{\Sigma_{00} + \Sigma_{11}}{2}\,.$$

Moreover, equality implies that F_0 and F_1 have the same cardinality.

Proof. We will in fact prove the stronger fact:

$$\Sigma_{00} - 2\Sigma_{01} + \Sigma_{11} \geqslant \big(|F_0| - |F_1|\big)^2 \min_{\alpha\in F} x_\alpha\,, \tag{4.6}$$

where $|\cdot|$ denotes set cardinality. The proof is by induction on $|F|$. It clearly holds for $|F| = 1$. Fix some n and assume that (4.6) always holds when $|F| = n$. Take a set F with $|F| = n+1$, and take positive numbers x_α, $\alpha \in F$. We can assume that $F = \{1, \dots, n+1\}$ and that $x_1 \geqslant \cdots \geqslant x_{n+1}$. Take any partition $F = F_0 \sqcup F_1$. Without loss of generality, assume that $n+1 \in F_0$. Apply the induction hypothesis to $F' = \{1, \dots, n\}$, obtaining

$$\Sigma'_{00} - 2\Sigma'_{01} + \Sigma'_{11} \geqslant \big(|F_0| - 1 - |F_1|\big)^2 x_n.$$

We have

$$\Sigma_{00} = \Sigma'_{00} + \big(2|F_0| - 1\big)x_{n+1}\,, \quad \Sigma_{01} = \Sigma'_{01} + |F_1| x_{n+1}\,, \quad \text{and} \quad \Sigma_{11} = \Sigma'_{11}\,,$$

so (4.6) follows. ☐

If $\mathtt{R}$ is an island or the world, let $\operatorname{pop}_1\mathtt{R}$ denote the banner 1 population on $\mathtt{R}$, that is, the sum of the populations of the cities in $\mathtt{R}$ with banner 1.

Let us give the following useful consequence of Lemma 4.13:

Lemma 4.15. $\operatorname{acyc}\mathrm{Ad}_A = \operatorname{pop}_1\mathtt{W}$.

Proof. By Proposition 3.3, $\operatorname{acyc}\operatorname{Ad}_A$ is the maximum of the geometric multiplicities of the eigenvalues of Ad_A. Those eigenvalues are the banners β, and the geometric multiplicity of each β is the worldwide total population with banner β. Thus, to prove the lemma we have to show that banner 1 has biggest worldwide population.

Let β be a banner. Then, using Lemma 4.13,

$$\sum_{\substack{\mathtt{C}\text{ is a city}\\ \text{with banner }\beta}} \operatorname{pop}\mathtt{C} \leqslant \frac{1}{2}\sum_{\substack{\mathtt{C}\text{ is a city}\\ \text{with banner }\beta}} \operatorname{pop}\pi_{\mathrm{r}}(\mathtt{C}) + \frac{1}{2}\sum_{\substack{\mathtt{C}\text{ is a city}\\ \text{with banner }\beta}} \operatorname{pop}\pi_{\mathrm{c}}(\mathtt{C})\,.$$

Since no two cities in the same row (resp. column) can have the same banner, the restriction of π_{r} (resp. π_{c}) to the set of cities with banner β is a one-to-one map. This allows us to conclude. □

Remark 4.16. The *Jordan type* of a matrix $A \in \operatorname{Mat}_{d\times d}(\mathbb{C})$ consists on the following data:

1. The number of different eigenvalues.
2. For each eigenvalue, the number of Jordan blocks and their sizes.

It follows from Lemma 4.15 that these data is sufficient to determine $\operatorname{acyc}\operatorname{Ad}_A$. (Of course, one can easily write down a formula; see e.g. [Ga, p. 222] or [Ar, p. 241].)

4.6. **Rigidity estimate for islands.**

Lemma 4.17. *For any island* $\mathtt{I}$,

$$\operatorname{rig}_+(\operatorname{Ad}_A|\mathtt{I}^\square) \leqslant \frac{\operatorname{pop}_1\pi_{\mathrm{r}}(\mathtt{I}) + \operatorname{pop}_1\pi_{\mathrm{c}}(\mathtt{I})}{2}\,.$$

In order to prove this lemma, it is convenient to consider separately the cases of non-equatorial and equatorial islands.

Proof of Lemma 4.17 when $\mathtt{I}$ *is non-equatorial.* For each banner β in $\mathtt{I}$, let n_β (resp. s_β) be the maximum of $\operatorname{rig}_+(\operatorname{Ad}_A|\mathtt{C}^\square)$ over the northern (resp. southern) cities $\mathtt{C}$ in $\mathtt{I}$ with banner β. For each city $\mathtt{C}$ with banner β, choose matrices $X_{\mathtt{C},1}$, …, $X_{\mathtt{C},n_\beta+s_\beta} \in \mathtt{C}^\square$ such that:

- $\Lambda_{\mathtt{C}} := \operatorname{sorb}_{\operatorname{Ad}_A}(X_{\mathtt{C},1},\dots,X_{\mathtt{C},m})$ is a transitive subspace of $\mathtt{C}^\square$;
- if $\mathtt{C}$ is southern then $X_1 = X_2 = \cdots = X_{n_\beta} = 0$;
- if $\mathtt{C}$ is northern then $X_{n_\beta+1} = \cdots = X_{n_\beta+s_\beta} = 0$.

Also, let $X_{\mathtt{C},j} = 0$ for $j > n_\beta + s_\beta$.

Next, define

$$Y_{\beta,j} = \sum_{\substack{\mathtt{C}\text{ is a city}\\ \text{of }\mathtt{I}\text{ with banner }\beta}} X_{\mathtt{C},j} \tag{4.7}$$

and

$$Z_j = \sum_{\beta\text{ banner on }\mathtt{I}} Y_{\beta,j} \tag{4.8}$$

Consider the space

$$\Delta = \operatorname{sorb}_{\operatorname{Ad}_A}(Z_1,\dots,Z_m), \quad \text{where} \quad m = \max_{\beta\text{ banner on }\mathtt{I}}(n_\beta+s_\beta)$$

It follows from Lemma 3.2 that

$$\Delta = \operatorname{sorb}_{\operatorname{Ad}_A}\big\{Y_{\beta,j};\ \beta\text{ is a banner, } 1 \leqslant j \leqslant n_\beta+s_\beta\big\}.$$

Recall notation (3.4). We claim that

$$\Lambda_{\mathtt{C}} \subset \Delta^{[\mathtt{C}]}. \tag{4.9}$$

Indeed, given $M \in \Lambda_{\mathtt{C}}$, write $M = \sum_j f_j(\operatorname{Ad}_A)X_{\mathtt{C},j}$, where the f_j's are polynomials and $f_j \equiv 0$ whenever $X_{\mathtt{C},j} = 0$. Consider $N = \sum_j f_j(\operatorname{Ad}_A)Y_{\beta,j}$, where β is

the banner of $\mathtt{C}$. Then it follows from Lemma 4.2 (part 2) that $N \in \Delta^{[\mathtt{C}]}$. This shows (4.9). So, by Lemma 3.11, Δ is a transitive subspace of $\mathtt{I}^\square$, showing that $\mathrm{rig}_+(\mathrm{Ad}_A|\mathtt{I}^\square) \leqslant m$.

To complete the proof of the lemma in the non-equatorial case, we will show that

$$m \leqslant \frac{\mathrm{pop}_1\,\pi_\mathrm{r}(\mathtt{I}) + \mathrm{pop}_1\,\pi_\mathrm{c}(\mathtt{I})}{2}\,. \tag{4.10}$$

Let β be the banner for which $n_\beta + s_\beta$ attains the maximum m. If $n_\beta > 0$, let $\mathtt{C}_N$ be a northern city within $\mathtt{I}$ with banner β and $\mathrm{rig}_+(\mathrm{Ad}_A|\mathtt{C}_N^\square) = n_\beta$. If $s_\beta > 0$, let $\mathtt{C}_S$ be a southern city within $\mathtt{I}$ with banner β and $\mathrm{rig}_+(\mathrm{Ad}_A|\mathtt{C}_S^\square) = s_\beta$. Assume for the moment that both cities exist. Let $\mathtt{C}_1$, $\mathtt{C}_2$, $\mathtt{C}_3$, $\mathtt{C}_4$ be projected equatorial cities as in Fig. 3.

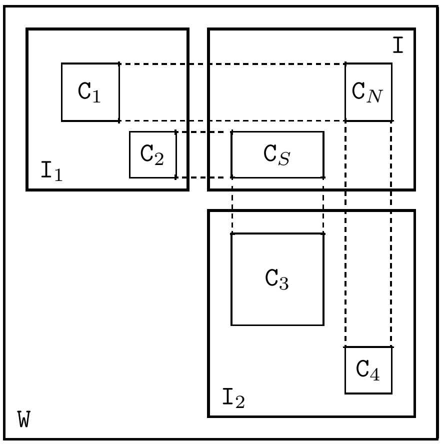

FIGURE 3. $\mathtt{C}_1 = \pi_\mathrm{r}(\mathtt{C}_N)$, $\mathtt{C}_2 = \pi_\mathrm{r}(\mathtt{C}_S)$, $\mathtt{C}_3 = \pi_\mathrm{c}(\mathtt{C}_S)$, $\mathtt{C}_4 = \pi_\mathrm{c}(\mathtt{C}_N)$.

Then

$$\begin{aligned} m = \mathrm{rig}_+(\mathrm{Ad}_A|\mathtt{C}_N) + \mathrm{rig}_+(\mathrm{Ad}_A|\mathtt{C}_S) &\overset{\text{(i)}}{\leqslant} \mathrm{pop}\,\mathtt{C}_N + \mathrm{pop}\,\mathtt{C}_S \\ &\overset{\text{(ii)}}{\leqslant} \tfrac{1}{2}\big(\mathrm{pop}\,\mathtt{C}_1 + \cdots + \mathrm{pop}\,\mathtt{C}_4\big) \leqslant \tfrac{1}{2}\big(\mathrm{pop}_1\,\mathtt{I}_1 + \mathrm{pop}_1\,\mathtt{I}_2\big), \end{aligned}$$

where (i) and (ii) follow respectively from Lemmas 4.12 and 4.13. This proves (4.10) in this case. If there is no southern city or no northern city within $\mathtt{I}$ with banner 1 then the proof of (4.10) is easier.

So the lemma is proved for non-equatorial $\mathtt{I}$. □

We now consider equatorial islands. There is an exceptional kind of island for which the proof of the rigidity estimate has to follow a different strategy. An island is called *exotic* if it has only the banners 1 and -1 (so it is equatorial and has 4 cities), each city has a single district, and all districts have the same population.

Proof of Lemma 4.17 when $\mathtt{I}$ *is equatorial non-exotic.* As in the previous case, let n_β (resp. s_β) be the maximum of $\mathrm{rig}_+(\mathrm{Ad}_A|\mathtt{C}^\square)$ over the northern (resp. southern) cities $\mathtt{C}$ in $\mathtt{I}$ with banner β.

We claim that

$$n_\beta + s_\beta < \mathrm{pop}_1\,\mathtt{I} \quad \text{for all banners } \beta \neq 1 \text{ in } \mathtt{I}. \tag{4.11}$$

Let us postpone the proof of this inequality and see how to conclude.

Let $M = \mathrm{pop}_1\,\mathtt{I}$. In view of Lemma 4.12 and relation (4.11), for each island $\mathtt{C}$ we can take matrices $X_{\mathtt{C},1}, \ldots, X_{\mathtt{C},M} \in \mathtt{C}^\square$ such that:

- $\Lambda_{\mathtt{C}} := \mathrm{sorb}_{\mathrm{Ad}_A}(X_{\mathtt{C},1}, \dots, X_{\mathtt{C},M})$ is a transitive subspace of $\mathtt{C}^{\square}$;
- $X_{\mathtt{C},M} = 0$ if $\mathtt{C}$ is non-equatorial;
- $X_{\mathtt{C},M}$ is the identity in $\mathtt{C}^{\square}$ if $\mathtt{C}$ is equatorial.

Then define matrices Z_j as before: by (4.7) and (4.8). Here we have that Z_M is the identity matrix in $\mathtt{I}^{\square}$. As before, $\mathrm{sorb}_{\mathrm{Ad}_A}(Z_1, \dots, Z_M)$ is a transitive subspace of $\mathtt{I}^{\square}$. Hence $\mathrm{rig}_+(\mathrm{Ad}_A|\mathtt{I}^{\square}) \leqslant M = \mathrm{pop}_1\,\mathtt{I}$, as desired.

Now let us prove (4.11). Consider a banner $\beta \neq 1$ in $\mathtt{I}$. Let $\mathtt{C}_N$ (resp. $\mathtt{C}_S$) be a northern (resp. southern) city within $\mathtt{I}$ with banner β and of maximal population; assume for the moment that both cities exist. Let $\mathtt{C}_1$, $\mathtt{C}_2$, $\mathtt{C}_3$, $\mathtt{C}_4$ be projected equatorial cities as in Fig. 4.

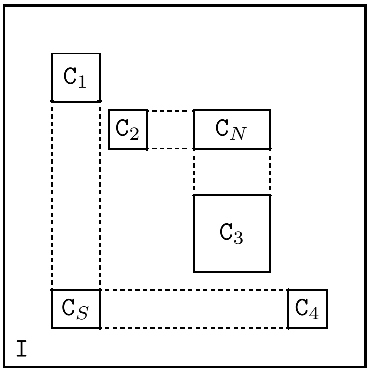

FIGURE 4. $\mathtt{C}_1 = \pi_{\mathrm{c}}(\mathtt{C}_S)$, $\mathtt{C}_2 = \pi_{\mathrm{r}}(\mathtt{C}_N)$, $\mathtt{C}_3 = \pi_{\mathrm{c}}(\mathtt{C}_N)$, $\mathtt{C}_4 = \pi_{\mathrm{r}}(\mathtt{C}_S)$. It is possible that $\mathtt{C}_1 = \mathtt{C}_2$ or $\mathtt{C}_3 = \mathtt{C}_4$.

Then

$$
\begin{aligned}
n_\beta + s_\beta &= \mathrm{rig}_+(\mathrm{Ad}_A|\mathtt{C}_N) + \mathrm{rig}_+(\mathrm{Ad}_A|\mathtt{C}_S) \\
&\leqslant \mathrm{pop}\,\mathtt{C}_N + \mathrm{pop}\,\mathtt{C}_S \qquad (4.12)\\
&\leqslant \tfrac{1}{2}\big(\mathrm{pop}\,\mathtt{C}_1 + \dots + \mathrm{pop}\,\mathtt{C}_4\big) \qquad (4.13)\\
&\leqslant \mathrm{pop}_1\,\mathtt{I}. \qquad (4.14)
\end{aligned}
$$

Inequality (4.12) follows from Lemma 4.12, inequality (4.13) follows from Lemma 4.13, and inequality (4.14) holds because the cities $\mathtt{C}_1$, …, $\mathtt{C}_4$ are equatorial, and any city can appear at most twice in this list. So

$$
n_\beta + s_\beta \leqslant \mathrm{pop}_1\,\mathtt{I}. \tag{4.15}
$$

In the case that there is no northern city or no southern city with banner β (i.e., n_β or s_β vanishes), a simpler argument shows that strict inequality holds in (4.15).

Now assume by contradiction that (4.11) does not hold. Then we must have equality in (4.15). By what we just saw, both cities $\mathtt{C}_N$ and $\mathtt{C}_S$ above exist. Then the inequalities in (4.12)–(4.14) become equalities. Since (4.14) is an equality, there must be exactly two equatorial cities in $\mathtt{I}$. So the non-equatorial banner β satisfies $\beta^{-1} = \beta$, that is, $\beta = -1$. Since (4.13) is an equality, it follows from Lemma 4.13 that both non-equatorial cities are district-square. So there is some ℓ such that all four cities in $\mathtt{I}$ have ℓ rows of districts and ℓ columns of districts. Since (4.12) is an equality, Lemma 4.12 implies that $\ell = 1$. That is, $\mathtt{I}$ is a exotic island, a situation which we excluded a priori. This contradiction proves (4.11) and Lemma 4.17 in the present case. $\square$

We now come to exotic islands. In all the previous cases, the transitive subspace we found had some vaguely Toeplitz form. For exotic islands, however, this strategy

is not efficient.[13] What we are going to do is to find a transitive space of vaguely Hankel form, namely the following:

$$\Lambda_k = \left\{ \begin{pmatrix} P & M \\ M & N \end{pmatrix} ;\ M,\ N,\ P \text{ are } k \times k \text{ matrices} \right\} . \tag{4.16}$$

Notice that $\Lambda_k = S_k \cdot \Gamma_k$, where

$$S_k = \begin{pmatrix} 0 & \mathrm{Id} \\ \mathrm{Id} & 0 \end{pmatrix} \quad \text{and} \quad \Gamma_k = \left\{ \begin{pmatrix} M & N \\ P & M \end{pmatrix} ;\ M,\ N,\ P \text{ are } k \times k \text{ matrices} \right\} .$$

Since Γ_k is a generalized Toeplitz space, it follows from Remark 2.3 that Λ_k is transitive.

Proof of Lemma 4.17 when $\mathtt{I}$ *is exotic.* If $\mathtt{I}$ is exotic then it has size $2k \times 2k$ for some k, and the operator $\mathrm{Ad}_A|\mathtt{I}^\square$ is given by $X \mapsto \mathrm{Ad}_L(X)$, where

$$L = \begin{pmatrix} D & 0 \\ 0 & -D \end{pmatrix} , \quad \text{and } D = D_k \text{ is the Jordan block (4.3).}$$

Let V be unique Ad_D-invariant subspace of $\mathrm{Mat}_{k\times k}(\mathbb{C})$ that has codimension 1 and does not contain the identity matrix (which exists by Lemma 4.3). Take matrices $X_1, \dots, X_k \in \mathrm{Mat}_{k\times k}(\mathbb{C})$ such that $X_1 = \mathrm{Id}$ and $V = \mathrm{sorb}_{\mathrm{Ad}_D}(X_2, \dots, X_k)$. Define $Y_1, \dots, Y_k \in \mathrm{Mat}_{2k\times 2k}(\mathbb{C})$ by

$$Y_1 = \begin{pmatrix} \mathrm{Id} & 0 \\ 0 & \mathrm{Id} \end{pmatrix} , \quad Y_j = \begin{pmatrix} X_j & 0 \\ 0 & 0 \end{pmatrix} \text{ for } 2 \leqslant j \leqslant k,$$

Then

$$\mathrm{sorb}_{\mathrm{Ad}_L}(Y_1, \dots, Y_k) = \left\{ \begin{pmatrix} x\mathrm{Id} + K & 0 \\ 0 & x\mathrm{Id} \end{pmatrix} ;\ x \in \mathbb{C},\ K \in V \right\} .$$

For $j = k+1, \dots, 2k$, define

$$Y_j = \begin{pmatrix} 0 & X_{j-k} \\ X_{j-k} & X_{j-k} \end{pmatrix} .$$

Then, by Lemma 3.2,

$$\mathrm{sorb}_{\mathrm{Ad}_L}(Y_{k+1}, \dots, Y_{2k}) = \left\{ \begin{pmatrix} 0 & M \\ M & N \end{pmatrix} ;\ M,\ N \in \mathrm{Mat}_{k\times k}(\mathbb{C}) \right\} .$$

Therefore $\mathrm{sorb}_{\mathrm{Ad}_L}(Y_1, \dots, Y_{2k})$ is the transitive space given by (4.16). Since Y_1 is the identity on $\mathtt{I}$, this shows that $\mathrm{rig}_+(\mathrm{Ad}_A|\mathtt{I}^\square) \leqslant 2k = \mathrm{pop}_1\,\mathtt{I}$, concluding the proof of Lemma 4.17. ☐

4.7. **The final rigidity estimate.** Let $c = c(A)$ be the number of equivalence classes mod T of eigenvalues of A.

Lemma 4.18. *If* $c < d$ *then*

$$\mathrm{rig}_+\,\mathrm{Ad}_A \leqslant \mathrm{pop}_1\,\mathtt{W} - c + 1 .$$

Proof. Let $m = \mathrm{pop}_1\,\mathtt{W} - c + 1$. For each island $\mathtt{I}$, let

$$r(\mathtt{I}) = \left\lfloor \tfrac{1}{2}(\mathrm{pop}_1\,\pi_{\mathrm{r}}(\mathtt{I}) + \mathrm{pop}_1\,\pi_{\mathrm{c}}(\mathtt{I})) \right\rfloor .$$

We claim that

$$r(\mathtt{I}) \leqslant \begin{cases} m & \text{if } \mathtt{I} \text{ is an equatorial island,} \\ m-1 & \text{if } \mathtt{I} \text{ is a non-equatorial island.} \end{cases} \tag{4.17}$$

Let us postpone the proof of this and see how to conclude the lemma.

In view of Lemma 4.17 and relation (4.17), for each island $\mathtt{I}$ we can take matrices $X_{\mathtt{I},1}, \dots, X_{\mathtt{I},m} \in \mathtt{I}^\square$ such that:

[13]For those who read Appendix A, notice that the simplest exotic island appears when A has a type 3 constraint; we have dealt with them in the proof of Proposition A.2.

- $\Lambda_{\mathtt{I}} := \mathrm{sorb}_{\mathrm{Ad}_A}(X_{\mathtt{I},1}, \dots, X_{\mathtt{I},m})$ is a transitive subspace of $\mathtt{I}^{\square}$;
- $X_{\mathtt{I},m} = 0$ if $\mathtt{I}$ is non-equatorial;
- $X_{\mathtt{I},m}$ is the identity in $\mathtt{I}^{\square}$ if $\mathtt{I}$ is equatorial.

Define matrices:

$$Y_{\alpha,j} = \sum_{\substack{\mathtt{I} \text{ is an island} \\ \text{with banner class } \alpha}} X_{\mathtt{I},j} \qquad (\alpha \text{ is a banner class, } 1 \leqslant j \leqslant m),$$

$$Z_j = \sum_{\alpha \text{ is a banner class}} Y_{\alpha,j} \qquad (1 \leqslant j \leqslant m).$$

So Z_m is the $d \times d$ identity matrix. Consider the space

$$\Delta = \mathrm{sorb}_{\mathrm{Ad}_A}(Z_1, \dots, Z_m).$$

It follows from Lemma 3.2 that

$$\Delta = \mathrm{sorb}_{\mathrm{Ad}_A} \big\{Y_{\alpha,j};\ \alpha \text{ is a banner class, } 1 \leqslant j \leqslant m\big\}.$$

We claim that every island $\mathtt{I}$,

$$\Lambda_{\mathtt{I}} \subset \Delta^{[\mathtt{I}]}. \tag{4.18}$$

Indeed, if $M \in \mathtt{I}$ then we can write $M = \sum_j f_j(\mathrm{Ad}_A) X_{\mathtt{I},j}$, where the f_j's are polynomials. Consider $N = \sum_j f_j(\mathrm{Ad}_A) Y_{\alpha,j}$, where α is the banner class of $\mathtt{I}$. It follows Lemma 4.2 (part 1) that $N \in \Delta^{[\mathtt{I}]}$. This proves (4.18). So, by Lemma 3.11, Δ is a transitive subspace of $\mathrm{Mat}_{d\times d}(\mathbb{C})$, showing that $\mathrm{rig}_+ \mathrm{Ad}_A \leqslant m$.

To conclude the proof we have to show estimate (4.17). First consider a equatorial island $\mathtt{I}$. Since there are c equatorial islands, and each of them has a positive banner 1 population, we conclude that $r(\mathtt{I}) \leqslant m$, as claimed.

Now take a non-equatorial $\mathtt{I}$. Applying what we just proved for the equatorial islands $\pi_{\mathrm{r}}(\mathtt{I})$ and $\pi_{\mathrm{c}}(\mathtt{I})$, we conclude that $r(\mathtt{I}) \leqslant m$. Now assume that (4.17) does not hold for $\mathtt{I}$, that is, $r(\mathtt{I}) = m$. Then

$$\mathrm{pop}_1 \pi_{\mathrm{r}}(\mathtt{I}) = \mathrm{pop}_1 \pi_{\mathrm{c}}(\mathtt{I}) = m = \mathrm{pop}_1 \mathtt{W} - c + 1.$$

Since $\mathrm{pop}_1 \mathtt{W} \geqslant \mathrm{pop}_1 \pi_{\mathrm{r}}(\mathtt{I}) + \mathrm{pop}_1 \pi_{\mathrm{c}}(\mathtt{I}) + c - 2$, we have $m = 1$ and $\mathrm{pop}_1 \mathtt{W} = c$. This means that $\mathrm{pop}_1 \tilde{\mathtt{I}} = 1$ for all equatorial islands $\tilde{\mathtt{I}}$, which is only possible if $c = d$. However, this case was excluded by hypothesis.

This proves (4.17) and hence Lemma 4.18. □

Example 4.19. If A is the matrix of Example 4.1 then Lemma 4.18 gives the estimate $\mathrm{rig}_+ \mathrm{Ad}_A \leqslant 28$. A more careful analysis (going through the proofs of the lemmas) would give $\mathrm{rig}_+ \mathrm{Ad}_A \leqslant 7$ (see Example 4.11).

Proof of part 2 of Theorem 3.7. Apply Lemmas 4.15 and 4.18. □

5. Proof of the hard part of the codimension m theorem

We showed in Proposition 2.11 that $\mathrm{codim}\, \mathcal{P}_m^{(\mathbb{K})} \leqslant m$. In this section, we will prove the reverse inequalities. More precisely, we will first prove Theorem 1.9 and then deduce Theorem 1.8 from it.

5.1. **Preliminaries on elementary algebraic geometry.**

5.1.1. *Quasiprojective varieties.* An algebraic subset of $\mathbb{C}^n$ is also called an *affine variety.* A *projective variety* is a subset of $\mathbb{CP}^n$ that can be expressed as the zero set of a family of homogeneous polynomials in $n+1$ variables. The *Zariski topology* on an (affine or projective) variety X is the topology whose closed sets are the (affine or projective) subvarieties of X.

An open subset U of a projective variety X is called a *quasiprojective variety.* We consider in U the induced Zariski topology. The affine space $\mathbb{C}^n$ can be identified with a quasiprojective variety. namely its image under the embedding $(z_1, \dots, z_n) \mapsto (1 : z_1 : \dots : z_n)$.

If X and Y are quasi-projective varieties then the product $X \times Y$ can be identified with a quasiprojective variety, namely its image under the Segre embedding; see [Sh, § 5.1].

The following is an important and very useful property of projective varieties. (See [Sh, p. 58] for a proof).

Proposition 5.1. *If X is a projective variety and Y is a quasiprojective variety then the projection $p\colon X \times Y \to Y$ takes Zariski closed sets to Zariski closed sets.*

A quasiprojective variety is called *irreducible* if it cannot be written as a nontrivial union of two quasiprojective varieties (that is, none contains the other).

5.1.2. *Dimension.* The dimension $\dim X$ of an irreducible quasiprojective variety X may be defined in various equivalent ways (see for instance [Ha, p. 133ff]). It will be sufficient for us to know that there exists an (intrinsically defined) subvariety Y of the *singular points of* X such that in a neighborhood of each point of $X \smallsetminus Y$, the set X is a complex submanifold of dimension (in the classical sense of differential geometry) $\dim X$; moreover, each irreducible component of Y has dimension strictly less than $\dim X$.

The dimension of a general quasiprojective variety is by definition the maximum of the dimensions of the irreducible components.

Remark 5.2. The dimension of a quasiprojective variety $U \subset \mathbb{CP}^n$ coincides with the dimension of its Zariski-closure in $\mathbb{CP}^n$ (see [Ha, p. 135]).

The following lemma is useful to estimate the codimension of an algebraic set X from information about the fibers of a certain projection $\pi\colon X \to Y$.[14]

Lemma 5.3. *Let Y be a quasiprojective variety. Let $X \subset Y \times \mathbb{CP}^n$ be a nonempty algebraically closed set. Let $\pi\colon X \to Y$ be the projection along $\mathbb{CP}^n$. Then:*

1. *For each $j \geqslant 0$, the set*
$$C_j = \{y \in Y;\ \operatorname{codim} \pi^{-1}(y) \leqslant j\}$$
is algebraically closed in Y.
2. *The dimension of X is given in terms of the dimensions of the C_j's by:*

$$\operatorname{codim} X = \min_{j;\ C_j \neq \varnothing} \left(j + \operatorname{codim} C_j\right). \tag{5.1}$$

The lemma is a consequence of standard theorems in algebraic geometry but for the reader's convenience let us spell out the details.

Proof of Lemma 5.3. In what follows, all topologies are of course Zariski. We will prove the equivalent "dual form" of the lemma, namely, that the sets
$$Y_k = \left\{y \in \pi(X);\ \dim \pi^{-1}(y) \geqslant k\right\}$$
are algebraically closed in Y, and
$$\dim X = \max_{k;\ Y_k \neq \varnothing} \big(k + \dim Y_k\big). \tag{5.2}$$

[14] A slightly similar result is [SW, Prop. 16].

First, the sets $X_k = \{x \in X;\ \dim \pi^{-1}(\pi(x)) \geqslant k\}$ are closed. (see [Ha, Thrm. 11.12]). So, by Proposition 5.1, $Y_k = \pi(X_k)$ is closed.

For each k with $X_k \neq \varnothing$, let $X_{k,i}$ indicate the irreducible components of X_k. Let

$$\mu(k,i) = \min_{x \in X_{k,i}} \dim \pi^{-1}(\pi(x)) .$$

Then, by [Ha, Thrm. 11.12] (and Remark 5.2),

$$\dim X_{k,i} = \mu(k,i) + \dim \pi(X_{k,i}) .$$

By definition, $\mu(k,i) \geqslant k$; moreover equality holds unless $X_{k,i} \subset X_{k+1}$. So

$$X_{k,i} \not\subset X_{k+1} \ \Rightarrow\ \dim X_{k,i} = k + \dim \pi(X_{k,i}) \leqslant k + \dim Y_k .$$

Since $X = \bigcup_{X_{k,i} \not\subset X_{k+1}} X_{k,i}$, this proves the $\leqslant$ inequality in (5.2).

To prove the converse inequality, fix any k with $Y_k \neq \varnothing$. Find i such that $\dim \pi(X_{k,i}) = \dim Y_k$. Then

$$\dim X \geqslant \dim X_{k,i} = \mu(k,i) + \dim Y_k \geqslant k + \dim Y_k.$$

This proves (5.2) and hence the lemma. □

Remark 5.4. Lemma 5.3 works with the same statement if $\mathbb{C}\mathrm{P}^n$ is replaced by $\mathbb{C}^{n+1}$, provided one assumes that $X \subset Y \times \mathbb{C}^{n+1}$ is homogeneous in the second factor (i.e., $(y,z) \in X$ implies $(y,tz) \in X$ for every $t \in \mathbb{C}$). Indeed, this follows from the fact that the projection $\mathbb{C}^{n+1} \smallsetminus \{0\} \to \mathbb{C}\mathrm{P}^n$ preserves codimension of homogeneous sets.

5.1.3. *Dimension estimates for sets of vector subspaces.* If $M \in \mathrm{Mat}_{n\times m}(\mathbb{K})$, let $\mathrm{col}\, M \subset \mathbb{K}^n$ denote the column space of M. A set $X \subset \mathrm{Mat}_{n\times m}(\mathbb{K})$ is called *column-invariant* if

$$\left.\begin{array}{c} M \in X \\ N \in \mathrm{Mat}_{n\times m}(\mathbb{K}) \\ \mathrm{col}\, M = \mathrm{col}\, N \end{array}\right\} \ \Rightarrow\ N \in X.$$

So a column-invariant set X is characterized by its set of column spaces. We enlarge the latter set by including also subspaces, thus defining:

$$\llbracket X \rrbracket := \big\{ E \text{ subspace of } \mathbb{K}^n;\ E \subset \mathrm{col}\, M \text{ for some } M \in X \big\}. \tag{5.3}$$

Then we have:

Theorem 5.5. *Let $X \subset \mathrm{Mat}_{n\times m}(\mathbb{C})$ be an algebraically closed, column-invariant set. Suppose E is a vector subspace of $\mathbb{C}^n$ that does not belong to $\llbracket X \rrbracket$. Then*

$$\mathrm{codim}\, X \geqslant m + 1 - \dim E .$$

Theorem 5.5 follows without difficulty from intersection theory of the grassmannians ("Schubert calculus"). The proof is given in [BG1].[15]

5.1.4. *The real part of an algebraic set.* Let X be an algebraically closed subset of $\mathbb{C}^n$. The *real part* of X is defined as $X \cap \mathbb{R}^n$. This is an algebraically closed subset of $\mathbb{R}^n$. Indeed, generators of the corresponding ideal $f_1, \dots, f_k$ in $\mathbb{C}[T_1, \dots, T_n]$ can be replaced by the corresponding real and imaginary parts polynomials.

As in the complex case, there are many equivalent definitions of dimensions of real algebraic or semialgebraic sets, in algebraic geometry. We just point out that a real algebraic or semialgebraic set admits a stratification into real manifolds such that the maximal differential geometry dimension of the strata coincides with the algebraic geometry dimension (see [BCR, p. 50]).

Proposition 5.6. *If X is an algebraically closed subset of $\mathbb{C}^n$ then $\dim_{\mathbb{R}}(X \cap \mathbb{R}^n) \leqslant \dim_{\mathbb{C}} X$.*

[15] Theorem 5.5 has applications apart from the one given in the present paper. The proof given in [BG1] will be incorporated in [BG2].

We refer the reader to [Wh] for the definition of the map rnk which to a real (resp. complex) variety V and to point $p \in V$ associates the *real* (resp. *complex*) *rank of* V *at* p. In the same paper, the author shows the two following results:

- for any (real or complex) variety V, for any point $p \in V$ the rank $\mathrm{rnk}_p(V)$ is greater than the codimension of V.
- If V is real (resp. complex) there is a point $p \in V$ such that the real (resp. complex) rank satisfies $\mathrm{rnk}_p(V) = \mathrm{codim}_{\mathbb{R}}(V)$ (resp. $\mathrm{rnk}_p(V) = \mathrm{codim}_{\mathbb{C}}(V)$).
- Given a real variety $V \subset \mathbb{R}^n$, there is a unique smallest complex variety $V^* \subset \mathbb{C}^n$ containing V (in particular, V is the real part of V^*). Then we have $\mathrm{rnk}_p(V^*) = \mathrm{rnk}_p(V)$.

Proof of Proposition 5.6. Let V be the real variety $X \cap \mathbb{R}^n$. Let $p \in V$ such that $\mathrm{rnk}_p(V) = \mathrm{codim}_{\mathbb{R}}(V)$. Consider the unique smallest complex variety then $\mathrm{rnk}_p(V^*) = \mathrm{rnk}_p(V)$. In particular $\mathrm{codim}_{\mathbb{R}}(V^*) \geqslant \mathrm{codim}_{\mathbb{C}}(V^*)$. Since $V^* \subset X$, the proposition follows. ☐

5.2. **Rigidity and the dimension of the poor fibers.** For simplicity of notation, let us write $\mathcal{P}_m = \mathcal{P}_m^{(\mathbb{C})}$. Also, for $A \in \mathrm{GL}(d, \mathbb{C})$, write:

$$r(A) := \mathrm{rig}_+ \mathrm{Ad}_A - 1\,.$$

We decompose the set $\mathcal{P}_m$ of poor data in fibers:

$$\mathcal{P}_m = \bigcup_{A \in \mathrm{GL}(d,\mathbb{C})} \{A\} \times \mathcal{P}_m(A), \quad \text{where} \quad \mathcal{P}_m(A) \subset \mathfrak{gl}(d, \mathbb{C})^m\,. \tag{5.4}$$

Lemma 5.7. *For any* $A \in \mathrm{GL}(d, \mathbb{C})$*, the codimension of* $\mathcal{P}_m(A)$ *in* $\mathfrak{gl}(d, \mathbb{C})^m$ *is at least* $m + 1 - r(A)$*.*

The lemma follows easily from Theorem 5.5 above:

Proof. Fix $A \in \mathrm{GL}(d, \mathbb{C})$, and write $r = r(A)$. We can assume that $r \leqslant m$, otherwise there is nothing to prove. By definition, there exists a r-dimensional subspace $E \subset \mathfrak{gl}(d, \mathbb{C})^m$ such that $\mathrm{sorb}_{\mathrm{Ad}_A}(\mathrm{Id} \vee E)$ is transitive. Identify $\mathfrak{gl}(d, \mathbb{C})$ with $\mathbb{C}^{d^2}$ and thus regard $\mathcal{P}_m(A)$ as a subset of $\mathrm{Mat}_{d^2 \times m}(\mathbb{C})$. Since the set $\mathcal{P}_m$ is algebraically closed and saturated (recall § 2.3), the fiber $\mathcal{P}_m(A)$ is algebraically closed and column-invariant, as required by Theorem 5.5. In the notation (5.3), we have $E \notin [\![\mathcal{P}_m(A)]\!]$. So applying Theorem 5.5, the lemma is proved. ☐

5.3. **How rare is high rigidity?** For simplicity of notation, let us write:

$$a(A) := \mathrm{acyc}\,\mathrm{Ad}_A \quad \text{for } A \in \mathrm{GL}(d, \mathbb{C}).$$

So Theorem 3.7 says that $r(A) \leqslant a(A) - c(A)$ provided $c(A) < d$.

Lemma 5.8. *For any integer* $k \geqslant 1$*, the set*

$$M_k = \big\{A \in \mathrm{GL}(d, \mathbb{C});\ r(A) \geqslant k\big\};$$

is algebraically closed in $\mathrm{GL}(d, \mathbb{C})$*; moreover if* $M_k \neq \varnothing$ *then*

$$\mathrm{codim}\, M_k \begin{cases} = 0 & \text{if } k = 1, \\ \geqslant k & \text{if } k \geqslant 2. \end{cases}$$

Lemma 5.8 is basically a consequence of Theorem 3.7, using the following construction:

Lemma 5.9. *There is a family* $\mathcal{G}(A)$ *of subsets of* $\mathrm{GL}(d, \mathbb{C})$*, indexed by* $A \in \mathrm{GL}(d, \mathbb{C})$*, such that the following properties hold:*

- *Each* $\mathcal{G}(A)$ *contains* A*.*
- *Each* $\mathcal{G}(A)$ *is an immersed manifold of codimension* $a(A) - c(A)$*.*

- *There are only countably many different sets $\mathcal{G}(A)$.*

The *informal* proof of the lemma goes as follows: For each $A \in \mathrm{GL}(d, \mathbb{C})$, let $\mathcal{G}(A)$ be the set of matrices that have the same Jordan type as A (as defined in Remark 4.16), and (at least) the same mod T relations between the eigenvalues. Then $\mathcal{G}(A)$ contains the conjugacy class of A, which by Remark 3.4 has codimension $a(A)$. We can also move the eigenvalues (keeping the mod T relations); this gives $c(A)$ extra degrees of freedom, so the codimension of $\mathcal{G}(A)$ is $a(A) - c(A)$. Since there are only finitely many Jordan types of $d \times d$ matrices, and only countably many mod T relations, there are only countably many different sets $\mathcal{G}(A)$. A formal proof of Lemma 5.9 follows:

Proof. First suppose that $A \in \mathrm{GL}(d, \mathbb{C})$ is a matrix in Jordan form:

$$A = \begin{pmatrix} B_{t_1}(\lambda_1) & & \\ & \ddots & \\ & & B_{t_n}(\lambda_n) \end{pmatrix}, \text{ where } B_t(\lambda) := \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \in \mathrm{Mat}_{t\times t}(\mathbb{C}).$$

Let $c = c(A)$; by the definition (3.2), we can choose numbers

$$\mu_1, \dots, \mu_c \in \mathbb{C}_*, \quad \theta_1, \dots, \theta_n \in T, \quad k_1, \dots, k_n \in \{1, \dots, c\}$$

such that $\lambda_i = \theta_i \mu_{k_i}$ for each $i = 1, \dots, n$. Let U be the subset of $(y_1, \dots, y_c) \in \mathbb{C}^c$ such that

$$y_k \neq 0 \quad \text{for each } k, \tag{5.5}$$

$$i \neq j \Rightarrow \theta_i y_{k_i} \neq \theta_j y_{k_j}\,. \tag{5.6}$$

Define a map $\Phi\colon U \to \mathrm{GL}(d, \mathbb{C})$ by:

$$\Phi(y_1, \dots, y_c) = \begin{pmatrix} B_{n_1}(\theta_1 y_{k_1}) & & \\ & \ddots & \\ & & B_{n_k}(\theta_n y_{k_n}) \end{pmatrix}.$$

For every $y \in U$, condition (5.6) assures that $\Phi(y)$ has the same Jordan type as A, and therefore, by Remark 4.16, $a(\Phi(y)) = a(A)$.

We define the set $\mathcal{G}(A)$ as the image of the map $\Psi = \Psi_A\colon \mathrm{GL}(d, \mathbb{C}) \times U \to \mathrm{GL}(d, \mathbb{C})$ given by $\Psi(X, y) = \mathrm{Ad}_X(\Phi(y))$.

Let us check that property 5.9 holds. Let $\partial_1\Psi$ and $\partial_2\Psi$ denote the partial derivatives with respect to X and y, respectively. As we have seen in Remark 3.4, the rank of $\partial_1\Psi(X, y)$ is equal to $d^2 - a(\Phi(y)) = d^2 - a(A)$ for every (X, y). We claim that

$$(\partial_2\Psi(X, y))^{-1}(\text{image of } \partial_1\Psi(X, y)) = \{0\}; \tag{5.7}$$

To see this, consider the map $\Gamma\colon \mathrm{Mat}_{d\times d}(\mathbb{C}) \to \mathbb{C}^d$ that associates to each matrix the coefficients of its characteristic polynomial. Then $\partial_1(\Gamma \circ \Psi)(X, y) = 0$, while $\partial_2(\Gamma \circ \Psi)(0, 0)$ is one-to-one. So (5.7) follows. As a result, the rank of the derivative of Ψ is equal to $d^2 - a(A) + c(A)$ at every point. Therefore, by the Rank Theorem, the image of Ψ is an immersed manifold of codimension $a(A) - c(A)$.

For arbitrary $A \in \mathrm{GL}(d, \mathbb{C})$, we define $\mathcal{G}(A) = \mathcal{G}(A_0)$, where A_0 is the Jordan form of A. Each set $\mathcal{G}(A)$ depends only on the data n and $(t_i, \theta_i, k_i)_{i=1,\dots,n}$; therefore there are only countably many different sets $\mathcal{G}(A)$. ☐

Remark 5.10. It is not difficult to show that each $\mathcal{G}(A)$ is a actually a submanifold of $\mathrm{GL}(d, \mathbb{C})$, but we won't need this.

Proof of Lemma 5.8. If $k = 1$ then $M_1 = \mathrm{GL}(d, \mathbb{C})$ (since $d \geqslant 2$), so there is nothing to prove. Consider $k \geqslant 2$. We have already shown in § 2.3 that $\mathcal{P}_k$ is algebraic. Since $M_k = \{A \in \mathrm{GL}(d, \mathbb{C});\ \forall \hat{X} \in \mathfrak{gl}(d, \mathbb{C})^k,\ (A, \hat{X}) \in \mathcal{P}_k\}$, it is evident that M_k is algebraically closed as well. We are left to estimate its dimension.

Take a nonsingular point A_0 of M_k where the local dimension is maximal. Let D be the intersection of M_k with a small neighborhood of A_0; it is an embedded disk. Each $A \in D$ has $r(A) \geqslant 2$; therefore by (both parts of) Theorem 3.7, we have $a(A) - c(A) \geqslant r(A) \geqslant k$. So, in terms of the sets from Lemma 5.9,

$$D \subset \bigcup_{A \text{ s.t. } a(A)-c(A) \geqslant k} \mathcal{G}(A).$$

The right hand side is a countable union of immersed manifolds of codimension at least k. It follows (e.g. by Baire Theorem) that D (and hence M_k) has codimension at least k. □

5.4. **Proof of Theorem 1.9.** Now we apply Lemmas 5.7 and 5.8 to prove one of our major results:

Proof of Theorem 1.9. The set $\mathcal{P}_m \subset \mathrm{GL}(d, \mathbb{C}) \times [\mathfrak{gl}(d, \mathbb{C})]^m$ is homogeneous in the second factor. Using Lemma 5.3 together with Remark 5.4, we obtain that the sets

$$C_j = \big\{A \in \mathrm{GL}(d, \mathbb{C});\ \operatorname{codim} \mathcal{P}_m(A) \leqslant j\big\} \tag{5.8}$$

are algebraically closed in $\mathrm{GL}(d, \mathbb{C})$, and

$$\operatorname{codim} \mathcal{P}_m = \min_{j;\ C_j \neq \varnothing} \big(j + \operatorname{codim} C_j\big).$$

By Lemma 5.7, we have $C_j \subset M_{m+1-j}$. Therefore, by Lemma 5.8,

$$C_j \neq \varnothing \quad \Rightarrow \quad \operatorname{codim} C_j \begin{cases} \geqslant 0 & \text{if } j = m, \\ \geqslant m - j + 1 & \text{if } j \leqslant m - 1. \end{cases} \tag{5.9}$$

So $\operatorname{codim} \mathcal{P}_m \geqslant m$, as we wanted to show. □

The proof above only used that $\operatorname{codim} C_j \geqslant m - j$. On the other hand, using the full power of (5.9) we obtain:

Scholium 5.11. *The set of poor data in "fat fibers", namely*

$$\mathcal{F}_m := \big\{(A, B_1, \dots, B_m) \in \mathcal{P}_m^{(\mathbb{C})};\ \operatorname{codim} \mathcal{P}_m(A) \leqslant m - 1\big\},$$

has codimension at least $m + 1$ in $\mathrm{GL}(d, \mathbb{C}) \times [\mathfrak{gl}(d, \mathbb{C})]^m$.

Proof. The projection of $\mathcal{F}_m$ on $\mathrm{GL}(d, \mathbb{C})$ is C_{m-1}. Use Lemma 5.3 (together with Remark 5.4) and (5.9). □

5.5. **The real case.**

Proof of Theorem 1.8. The real part of $\mathcal{P}_m^{(\mathbb{C})}$ is a real algebraic set which, in view of Proposition 5.6, has codimension at least m. Recall from § 2.3 that this set contains the semialgebraic set $\mathcal{P}_m^{(\mathbb{R})}$, which therefore has codimension at least m. Since we already knew from Proposition 2.11 that $\mathcal{P}_m^{(\mathbb{R})} \leqslant m$, the theorem is proved. □

5.6. **Additional information.** Let us improve upon Scholium 5.11 and so prepare the ground for the proof of Theorem 1.2. This part is not necessary for the proof of Theorem 1.1.

Recall from § 2.3 the definition of saturated set.

Lemma 5.12. *There exists a saturated algebraically closed set $\mathcal{S}_m \subset \mathrm{GL}(d,\mathbb{C}) \times [\mathrm{Mat}_{d\times d}(\mathbb{C})]^m$ of codimension at least $m+1$ such that for all $(A, B_1, \dots, B_m) \in \mathcal{P}_m \smallsetminus \mathcal{S}_m$, the following properties hold:*

1. *A is unconstrained;*
2. *if $P \in \mathrm{GL}(d,\mathbb{C})$ is such that $P^{-1}AP$ is a diagonal matrix then there are indices i_0, $j_0 \in \{1, \dots, d\}$ with $i_0 \neq j_0$ such that for each $k \in \{1, \dots, m\}$, the (i_0, j_0) entry of the matrix $P^{-1}B_kP$ vanishes;*
3. *for each choice of P above, the off-diagonal vanishing entry position (i_0, j_0) is unique.*

Notice that each data in $\mathcal{P}_m \smallsetminus \mathcal{S}_m$, after a change of basis, satisfies precisely the hypotheses of Lemma 2.14.

In order to prove the lemma, we begin by checking algebraicity of the constraints:

Lemma 5.13. *The set $K \subset \mathrm{GL}(d,\mathbb{C})$ of constrained matrices is an algebraically closed subset of codimension* 1*.*

Proof. Multiply all constraints, obtaining a polynomial in the variables λ_1, …, λ_d. This polynomial is symmetric, and therefore (see e.g. [La, Thrm. IV.6.1]) can be written as a polynomial function of the elementary symmetric polynomials in the variables λ_1, …, λ_d. Now substitute each elementary symmetric polynomial in this expression by the corresponding coefficient of the characteristic polynomial of the matrix A. This gives a polynomial function on the entries of the matrix A that vanishes if and only if A is constrained. It is obvious that the corresponding algebraic set K has codimension 1. □

Now we check algebraicity of double vanishing:

Lemma 5.14. *There exists a saturated algebraically closed subset $\mathcal{D}$ of $\mathrm{GL}(d,\mathbb{C}) \times [\mathrm{Mat}_{d\times d}(\mathbb{C})]^m$ such that if $(A, B_1, \dots, B_m) \in \mathcal{D}$ and A has simple spectrum then property 2 from Lemma 5.12 is satisfied, but property 3 is not.*

Proof. First, consider the subset $X \subset [\mathrm{Mat}_{d\times d}(\mathbb{C})]^{1+m} \times (\mathbb{C}\mathrm{P}^{d-1})^2$ formed by tuples $(A, B_1, \dots, B_m, [v], [w])$ such that

$$[Av] = [v], \quad [A^*w] = [w], \quad w^*v = 0, \quad w^*B_kv = 0 \text{ for each } k = 1, \dots, m,$$

where v and w are regarded as column-vectors and the star denotes transposition. The set X is obviously algebraic; thus, by Proposition 5.1, so is its projection Y on $[\mathrm{Mat}_{d\times d}(\mathbb{C})]^{1+m}$.

Let A be a matrix with simple spectrum. Then $(A, B_1, \dots, B_m)$ belongs to Y if and only if property 2 from Lemma 5.12 is satisfied. In particular, the fiber of Y over A is a union of affine subspaces of $[\mathrm{Mat}_{d\times d}(\mathbb{C})]^m$. Intersections of those affine spaces correspond to points where the uniqueness property 3 is not satisfied. These points of intersection are singular points of Y. Conversely, it is clear that the variety Y is smooth at the points on the fiber over A where property 3 is satisfied.

So let Z be the (algebraically closed) set of singular points of Y. It is straightforward to see that the set Y is saturated. Recalling Remark 2.8 (part 1) and the fact that a group acting on a variety preserves singular points, we see that the set Z is saturated as well.

We define $\mathcal{D}$ as the set Z minus the tuples $(A, B_1, \dots, B_m)$ with $\det A = 0$. Then $\mathcal{D}$ has all the required properties. □

Proof of Lemma 5.12. For simplicity of writing we will omit the m subscripts.

Let $\pi : \mathcal{P} \to \mathrm{GL}(d, \mathbb{C})$ be the projection on the first matrix. Define

$$\mathcal{S} = \pi^{-1}(K) \cup (\mathcal{D} \cap \mathcal{P}),$$

where K and $\mathcal{D}$ come respectively from Lemmas 5.13 and 5.14. Then $\mathcal{S}$ is a saturated algebraically closed subset of $\mathcal{P}$. If $\mathbf{A} = (A, B_1, \dots, B_m) \in \mathcal{P} \smallsetminus \mathcal{S}$ then:

- $A \notin K$, which is property 1;
- since $\mathbf{A} \in \mathcal{P}$, it follows from Lemma 2.13 that $\mathbf{A}$ is conspicuously poor, and so property 2 holds;
- since $\mathbf{A} \notin \mathcal{D}$, property 3 also holds.

To complete the proof of the lemma, we need to show that $\operatorname{codim} \mathcal{S} \geqslant m + 1$. We will use the following inclusion:

$$\mathcal{S} \subset \mathcal{F} \cup \underbrace{\left(\pi^{-1}(K) \smallsetminus \mathcal{F}\right)}_{\mathcal{F}'} \cup \underbrace{\left((\mathcal{D} \cap \mathcal{P}) \smallsetminus \pi^{-1}(K)\right)}_{\mathcal{F}''} . \tag{5.10}$$

where $\mathcal{F}$ comes from Scholium 5.11. Recall that $\mathcal{F}$ equals $\pi^{-1}(C_{m-1})$, where C_j is given by (5.8), and it has codimension at least $m + 1$.

We apply Lemma 5.3 and Remark 5.4 to the set $\mathcal{F}' \subset Y' \times [\mathfrak{gl}(d, \mathbb{C})]^m$, where $Y' = \mathrm{GL}(d, \mathbb{C}) \smallsetminus C_{m-1}$. Since K has codimension at least 1 in Y', and the fibers of $\mathcal{F}'$ all have codimension at least m, we conclude that that $\operatorname{codim} \mathcal{F}' \geqslant m + 1$.

Next, we want to apply Lemma 5.3 and Remark 5.4 to the set $\mathcal{F}'' \subset Y'' \times [\mathfrak{gl}(d, \mathbb{C})]^m$, where $Y'' = \mathrm{GL}(d, \mathbb{C}) \smallsetminus K$. For each $A \in Y''$, it follows from Lemma 5.14 that the fiber of $\mathcal{F}''$ over A (which is the same as the fiber of $\mathcal{D}$ over A) has codimension $2m$ in $[\mathfrak{gl}(d, \mathbb{C})]^m$, corresponding to the $2m$ different matrix entries that must vanish. We conclude that $\operatorname{codim} \mathcal{F}'' \geqslant 2m$.

We have seen that each of the three sets on the right-hand side of (5.10) has codimension at least $m + 1$. So the same is true for $\mathcal{S}$, as we wanted to prove. $\square$

6. Proof of the main results

6.1. **Stratifications.** We first recall a few notions about stratifications. We refer the reader to [GWPL, Ma] for details and proofs.

Let Σ be a closed subset of a smooth (i.e., C^∞) manifold X. A *smooth stratification* of Σ is a filtration by closed subsets

$$\Sigma = \Sigma_n \supset \Sigma_{n-1} \supset \cdots \supset \Sigma_0$$

such that and for each i, the set $X_i = \Sigma_i \smallsetminus \Sigma_{i-1}$ (where $\Sigma_{-1} := \varnothing$) either is a smooth submanifold of M without boundary and of dimension i, or is empty. Each connected component of X_i is called a *stratum*. The *codimension* of a stratified space is the lowest codimension of strata. This does not depend on the choice of the stratification.

Note that, apart for discrete subsets $\Sigma \subset X$, if there is one smooth stratification, then there are infinitely many others. However, the subsets that we will be dealing with will be endowed with certain *canonical* stratifications:

Theorem 6.1 (Existence of canonical stratifications)**.** *Any algebraic set $\Sigma \subset \mathbb{C}^N$ admits a canonical smooth stratification, whose strata are complex submanifolds of $\mathbb{C}^N$. Any closed semialgebraic set $\Sigma \subset \mathbb{R}^N$ admits a canonical smooth stratification, whose strata are semialgebraic submanifolds of $\mathbb{R}^N$.*

In the case of an irreducible algebraic set $\Sigma \subset \mathbb{C}^n$, the canonical stratification can be obtained as follows: The connected components of the set of regular (i.e., non-singular) points form the higher-dimensional strata; then one decomposes the set of singular points of Σ into irreducible components and proceeds by induction.

In any case, those canonical stratifications are uniquely characterized by a certain minimality property. In particular, the canonical stratifications are equivariant under polynomial automorphisms of the ambient space.

Another important property of the canonical stratifications is that they satisfy the so-called *Whitney conditions*. We will not recall here those conditions, which would be rather technical; we will only write down some of their properties. A smooth stratification that satisfies the Whitney conditions is called a *Whitney stratification.*

Proposition 6.2 (Basic properties of Whitney stratifications)**.** *Let X, Y be smooth manifolds. Let*

$$\Sigma_n \supset \cdots \supset \Sigma_0 \tag{6.1}$$

be a filtration of a set $\Sigma \subset X$. Then:

1. *Being a Whitney stratification is a local property of a filtration: So if* (6.1) *is a Whitney stratification of Σ then $\Sigma_n \cap U \supset \cdots \supset \Sigma_0 \cap U$ is a Whitney stratification of $\Sigma \cap U$, and conversely if each point in Σ has an open neighborhood $U \subset X$ such that $\Sigma_n \cap U \supset \cdots \supset \Sigma_0 \cap U$ is a Whitney stratification of $\Sigma \cap U$ then* (6.1) *is a Whitney stratification of Σ.*
2. *If* (6.1) *is a Whitney stratification of Σ then $\Sigma_n \times Y \supset \cdots \supset \Sigma_0 \times Y$ is is a Whitney stratification of $\Sigma \times Y \subset X \times Y$.*
3. *If* (6.1) *is a Whitney stratification of Σ and $f\colon X \to Y$ is a smooth diffeomorphism then $f(\Sigma_n) \supset \cdots \supset f(\Sigma_0)$ is a Whitney stratification of $f(\Sigma) \subset Y$.*

Let us now discuss how stratifications behave with respect to transversality. Let $f\colon X \to Y$ be a C^1 map. Let $\Sigma = \Sigma_d \supset \cdots \supset \Sigma_0$ be a stratification of a closed subset Σ of Y. One says that f is *transverse* to that stratification if it is transverse to each of its strata. Transversality to a general stratification is not an open condition. However, we obtain openness if the stratification is Whitney:

Proposition 6.3 (Transversality is open)**.** *Let X, Y be C^∞ manifolds without boundary. Let $\Sigma = \Sigma_d \supset \cdots \supset \Sigma_0$ be a Whitney stratification of a closed subset of Y. Then the set $\mathcal{O} = \{f \in C^1(X,Y);\ f \pitchfork \Sigma\}$ is open in $C^1(X,Y)$ (with respect to the strong topology).*

Actually, only the first of the Whitney conditions is necessary here (use the (1)⇒(3) implication of Trotman's theorem [Tr]).

6.2. **Jets and jet transversality.** We recall the basic notions on jets and state the transversality theorems we will need; see [Hi] for details.

Let X, Y be smooth manifolds without boundary. If $1 \leqslant r < \infty$, an *r-jet from X to Y* is an equivalence class of pairs (x, f), where $x \in X$, f is a C^r map from a neighborhood of x to Y, and where (x, f) is equivalent to (x', f') if $x = x'$ and f and f' have same derivatives at x up to order r. We denote by $J^r(X,Y)$ the space of r-jets from X to Y. It is a smooth manifold.

For all $1 \leqslant s \leqslant \infty$, we denote by $C^s(X,Y)$ the space of C^s-maps from X to Y, endowed with the strong topology.

Given $1 \leqslant r < s \leqslant \infty$ and a map $g \in C^s(X,Y)$, the *r-jet extension* is the map $j^r g\colon X \to J^r(X,Y)$ that sends x to the equivalence class $j^r g(x)$ of (x, g). Then the mapping

$$j^r\colon C^s(X,Y) \to C^{s-r}\left(X, J^r(X,Y)\right)$$

is continuous.

Theorem 6.4 (Jet transversality)**.** *Let $1 \leqslant r < s \leqslant \infty$. Let X and Y be C^∞ manifolds without boundary. Let $W \subset J^r(X,Y)$ be a C^∞ submanifold without boundary.*

Then the C^s-maps $g\colon X \to Y$ for which the r-jet extension $j^r g$ is transverse to W form a residual subset of $C^s(X,Y)$.

Let us now show the following:

Proposition 6.5. *Let X, Y be C^∞-manifolds without boundary. Let $\Sigma \subset J^1(X,Y)$ be a Whitney stratified closed subset. Then $\{f \in C^2(X,Y);\ j^1 f \pitchfork \Sigma\}$ is C^2-open and C^∞-dense in $C^2(X,Y)$.*

Here, as in the introduction, we say that a subset of $C^2(X,Y)$ is C^∞-dense if its intersection with $C^r(X,Y)$ is C^r-dense, for every $r \geqslant 2$.

Proof of Proposition 6.5. By Proposition 6.3, the set $\{F\colon X \to j^1(X,Y);\ F \pitchfork \Sigma\}$ is open in $C^1\big(X, J^1(X,Y)\big)$. Hence the set $\mathcal{O} := \{f\colon X \to Y;\ j^1 f \pitchfork \Sigma\}$ is open in $C^2(X,Y)$.

Fix $r \geqslant 2$. Given a Whitney stratification $\Sigma_n \supset \cdots \supset \Sigma_0$ of Σ, let $Z_i = \Sigma_i \smallsetminus \Sigma_{i-1}$ be the corresponding decomposition into smooth submanifolds. By the jet transversality theorem (Theorem 6.4), each set $\mathcal{R}_i = \{f \in C^r(X,Y);\ j^1 f \pitchfork Z_i\}$ is residual. Thus $\mathcal{O} \cap C^r(X,Y) = \bigcap_i \mathcal{R}_i$ is C^r-dense. This concludes the proof. ☐

6.3. **Proof of the main result.** We now use Theorem 1.8 and the tools explained above to prove our main result. Before going into the proof itself, let us deal with a technical detail.

By Theorem 1.8, $\mathcal{P}_m^{(\mathbb{R})}$ is a closed semialgebraic subset of $\mathrm{GL}(d,\mathbb{R}) \times \mathfrak{gl}(d,\mathbb{R})^m$. Since Theorem 6.1 concerns semialgebraic subsets of affine space, we proceed as follows. First, enlarge $\mathcal{P}_m^{(\mathbb{R})}$ by including all $(A, B_1, \dots, B_m)$ with $\det A = 0$, thus obtaining a subset $\hat\Gamma$ of $[\mathrm{Mat}_{d\times d}(\mathbb{R})]^{1+m}$ which is also closed and semialgebraic. By Theorem 6.1, the set $\hat\Gamma$ admits a canonical Whitney stratification

$$\hat\Gamma = \hat\Gamma_n \supset \cdots \supset \hat\Gamma_0 .$$

Now we remove all $(A, B_1, \dots, B_m)$ with $\det A = 0$ from each $\hat\Gamma_i$, thus (by locality property 1 in Proposition 6.2) obtaining a Whitney stratification of codimension m:

$$\mathcal{P}_m^{(\mathbb{R})} = \Gamma_n \supset \cdots \supset \Gamma_0 . \tag{6.2}$$

(We may have $\Gamma_n = \Gamma_{n-1}$.) Since the stratification of $\hat\Gamma$ is canonical, the stratification (6.2) is invariant under polynomial automorphisms of the set $\mathrm{GL}(d,\mathbb{R}) \times \mathfrak{gl}(d,\mathbb{R})^m$ that preserve $\mathcal{P}_m^{(\mathbb{R})}$.

Proof of Theorem 1.1. Let $\mathcal{U}$ be a smooth manifold without boundary and of dimension m. Given local coordinates on an open set $U \subset \mathcal{U}$, the set of 1-jets from U to $\mathrm{GL}(d,\mathbb{R})$ may be identified with the set

$$U \times \mathrm{GL}(d,\mathbb{R}) \times \mathfrak{gl}(d,\mathbb{R})^m.$$

Indeed, a jet $\mathbf{J}$ represented by a pair (u, A) can be identified with the point

$$(u, A(u), B_1, \dots, B_m) \in U \times \mathrm{GL}(d,\mathbb{R}) \times \mathfrak{gl}(d,\mathbb{R})^m,$$

where $B_i \in \mathrm{Mat}_{d\times d}(\mathbb{R})$ is the normalized derivative of A at u, along the i^{th} coordinate. Let us say that the 1-jet $\mathbf{J}$ is *rich* if the data $\mathbf{A} = (A(u), B_1, \dots, B_m)$ is rich, or equivalently, if for sufficiently large N, the input $(u, \dots, u) \in \mathcal{U}^N$ is universally regular for the system (1.4). If the jet is not rich then it is called *poor*.

Define a filtration

$$\Sigma_n \supset \cdots \supset \Sigma_0 \tag{6.3}$$

of the set of poor jets from $\mathcal{U}$ to $\mathrm{GL}(d,\mathbb{R})$ as follows: a jet $\mathbf{J}$ represented as above in local coordinates by $(u, A(u), B_1, \dots, B_m)$ belongs to Σ_i if and only if

$(A(u), B_1, \dots, B_m)$ belongs to the set Γ_i in (6.2). We need to check that this definition does not depend on the choice of the local coordinates. Indeed, this follows from the fact that $\mathcal{P}_m^{(\mathbb{R})}$ is a saturated set (see § 2.3) using the invariance property of the stratification (6.2) explained above.

We claim that the filtration (6.3) is a Whitney stratification of codimension m. Indeed, the intersection of the filtration with the open subset $J^1(U, \mathrm{GL}(d,\mathbb{R}))$ of $J^1(\mathcal{U}, \mathrm{GL}(d,\mathbb{R}))$ is identified (through a smooth diffeomorphism) with the filtration

$$U \times \Gamma_n \supset \cdots \supset U \times \Gamma_0.$$

So the claim follows from Proposition 6.2.

Applying Proposition 6.5, we obtain a C^2-open C^∞-dense set $\mathcal{O} \subset C^2(\mathcal{U}, \mathrm{GL}(d,\mathbb{C}))$ formed by maps A that are transverse to the stratification (6.3) of the set of poor jets. Since the codimension of the stratification equals the dimension of the manifold $\mathcal{U}$, if $A \in \mathcal{O}$ then the points u for which $j^1 A(u)$ is poor form a 0-dimensional set. This proves Theorem 1.1. □

Remark 6.6. In the proof above, instead of working with the semialgebraic set $\mathcal{P}_m^{(\mathbb{R})}$, we could have worked equally well with the real part of $\mathcal{P}_m^{(\mathbb{C})}$, since it is an algebraic set containing $\mathcal{P}_m^{(\mathbb{R})}$ and has the same codimension.

6.4. **Proof of the addendum.**

Proof of Theorem 1.2. Consider the set $\mathcal{S}_m^{(\mathbb{C})}$ given by Lemma 5.12, and let $\mathcal{S}_m^{(\mathbb{R})}$ be its real part. This is an algebraically closed saturated subset of $\mathrm{GL}(d,\mathbb{R}) \times [\mathfrak{gl}(d,\mathbb{R})]^m$ which, by Proposition 5.6, has codimension at least $m+1$.

Consider the set $\tilde{\Gamma}$ of 1-jets $\mathbf{J} \in J^1(\mathcal{U}, \mathrm{GL}(d,\mathbb{C}))$ that have a local expression $(u, A(u), B_1, \dots, B_m)$ with $(A(u), B_1, \dots, B_m) \in \mathcal{S}_m^{(\mathbb{R})}$. This does not depend on the choice of the local coordinates, because $\mathcal{S}_m^{(\mathbb{R})}$ is saturated. By the same arguments as in the proof of Theorem 1.1, the set $\tilde{\Gamma}$ admits a Whitney stratification. Its codimension is at least $m+1$. Applying Proposition 6.5, we obtain a C^2-open C^∞-dense set $\tilde{\mathcal{O}} \subset C^2(\mathcal{U}, \mathrm{GL}(d,\mathbb{C}))$ formed by maps A that are transverse to the stratification.

Let $\mathcal{O}$ be the set provided by Theorem 1.1. and consider a map $A \in \mathcal{O} \cap \tilde{\mathcal{O}}$. Then whenever a jet $j^1 A(u)$ is poor, it does not belong to $\tilde{\Gamma}$. Recalling Lemma 5.12, we see that the local expression of $j^1 A(u)$ satisfies (after a change of basis) the hypotheses of Lemma 2.14. Therefore parts 1 and 2 of the theorem follow respectively from conclusions 1 and 2 of the lemma. □

Remark 6.7. The proof of Theorem 1.2 also gives more information about the 1-jets that appear generically for singular constant inputs $(u, \dots, u)$: the associated matrix data is conspicuously poor (see § 2.4), and the matrix $A(u)$ is unconstrained (see § 2.5).

Remark 6.8. Properties 1 and 2 in Theorem 1.2 are in fact dual to each other. If $\mathbf{A}$ is the data representing the 1-jet of A at u, and $\Lambda = \Lambda(\mathbf{A})$, then property 1 means that there is an unique direction $[v] \in \mathbb{R}\mathrm{P}^{d-1}$ such that $\Lambda \cdot v \neq \mathbb{C}^d$. Then property 2 means that there is an unique direction $[w] \in \mathbb{R}\mathrm{P}^{d-1}$ such that $\Lambda^* \cdot w \neq \mathbb{C}^d$, where Λ is the set of the transposes of the matrices in Λ. This fact can be proved easily using the dual characterization of Lemma 3.12.

Appendix A. The case of one-dimensional input

As we explained in § 1.4, this appendix contains a basically independent discussion of the case $\dim \mathcal{U} = 1$. The prerequisites are all contained in Section 2 and § 3.1. In order to avoid technicalities at this point, we will be sometimes informal, especially regarding questions of transversality.

Let us define the *canonical constraints* respectively of type 1, 2, 3, 4 as the following relations:

$$\lambda_1\lambda_3 = \lambda_2^2, \qquad \lambda_1\lambda_4 = \lambda_2\lambda_3, \qquad \lambda_1 = -\lambda_2, \qquad \lambda_1 = \lambda_2\,. \tag{A.1}$$

Recall from § 2.5 that a *constraint* between variables $\lambda_1, \ldots, \lambda_d$ is a relation that can be reduced to one of the four canonical constraints after a change of indices. Each constraint has a unique type.

Let us say that a matrix $A \in \mathrm{GL}(d,\mathbb{R})$ is (i)-*constrained*, for $1 \leqslant i \leqslant 4$ if:

- its eigenvalues, counted with multiplicity, satisfy exactly one elementary constraint, which is a type i constraint,
- if there is a type 4 constraint between the eigenvalues, then the matrix A is *not* diagonalizable.

Hence if a matrix A is not (i)-constrained for any $0 \leqslant i \leqslant 4$, then

- either A is unconstrained, i.e., its eigenvalues (with multiplicity) satisfy no constraint;
- or the eigenvalues of A satisfy at least two constraints;
- or A has a (multiple) eigenvalue corresponding to at least two Jordan blocks.

If either of the last two cases hold, we say that A is *multiconstrained.*

Proposition A.1. 1. *The complement of the set of unconstrained matrices has codimension* 1 *in* $\mathrm{GL}(d,\mathbb{R})$*.*

2. *The set of multiconstrained matrices has codimension* 2 *in* $\mathrm{GL}(d,\mathbb{R})$*.*

Informal proof. Matrices that are not unconstrained have at least one constraint on their eigenvalues, so the corresponding set has codimension 1.

Matrices that are very constrained either have at least two constraints on their eigenvalues, or have an eigenvalue of multiplicity 2 and are diagonalizable. In both cases, the corresponding set has codimension 2. □

Let us define *adapted basis* for matrices A that are not multiconstrained:

- If A is unconstrained then an adapted basis is a basis of eigenvectors.
- If A is (i)-constrained, for $i = 1$, 2, or 3 then an adapted basis is an (ordered) basis of eigenvectors such that the corresponding eigenvectors $\lambda_1, \ldots \lambda_d$ satisfy the canonical type i constraint.
- If A is (4)-constrained then an *adapted basis* for A is a basis in which A is written in the following *modified Jordan form*[16]:

$$\left(\begin{array}{cc|ccc} \lambda_1 & \lambda_1 & & & \\ 0 & \lambda_1 & & & \\ \hline & & \lambda_3 & & \\ & & & \ddots & \\ & & & & \lambda_d \end{array}\right).$$

Obviously, such adapted basis always exist.

If a matrix A is (i)-constrained then we say that a $d \times d$ matrix B is a *good match* for A, if there is an adapted basis for A in which it writes as $B = (b_{ij})$, where all nondiagonal entries b_{ij} are nonzero and if $b_{11} \neq b_{22}$, in the particular case where A is 3-constrained.

The usefulness of this definition is explained by the following Propositions A.2 and A.3[17]:

[16] The reason for using a modified Jordan form is that it makes the expression of Ad_A simpler, as we will see later.

[17] Actually, the definition of a good match matrix is stronger than necessary for Proposition A.2 to be true. But in order to avoid complications, we chose a condition that works for all types of constraints.

Proposition A.2. *If A is not multiconstrained and B is a good match for A then the pair (A,B) is rich.*

In other words, $\mathcal{P}_1^{(\mathbb{C})}$ is contained in the following set:

$$\begin{aligned}\text{(A.2)}\quad \mathcal{E} := \big\{(A,B)\in \mathrm{GL}(d,\mathbb{C})\times \mathfrak{gl}(d,\mathbb{C});\ &\text{either } A \text{ is multiconstrained} \\ &\text{or } A \text{ is not multiconstrained but } B \text{ is not a good match for } A\big\}.\end{aligned}$$

Proposition A.3. 1. *The set $\mathcal{E}$ has codimension 1.*
2. *The set $\{(A,B)\in\mathcal{E};\ A$ is not unconstrained$\}$ has codimension 2.*

Informal proof. This follows from Proposition A.1 and the fact that for each matrix A that is not multiconstrained, the set of B's that are not good matches for A has positive codimension in $\mathfrak{gl}(d,\mathbb{C})$. ☐

Theorem 1.9 in the case $m = 1$ follows from the propositions above. Therefore the other main results (Theorems 1.1, 1.2, 1.8 and C.1) in the $m = 1$ case also follow from the propositions. For any of these results, the propositions give extra information of practical value: with the explicit definition of the set $\mathcal{E}$ in (A.2), we know which 1-jets should be avoided in Theorem 1.1, for example. The discussion given in Appendix B also applies; it gives explicit conditions on the 2-jet extension of the map $A\colon \mathcal{U}\to \mathrm{GL}(d,\mathbb{R})$ that assure that A satisfies the conclusions of Theorems 1.1 and 1.2.

Proof of Proposition A.2. Let A and B satisfy the hypotheses. We need to show that the space $\Lambda(A,B)$ defined by (2.2) is a transitive subspace of $\mathfrak{gl}(d,\mathbb{C})$. Let $\Gamma := \mathrm{sorb}_{\mathrm{Ad}_A}(B)$, so that $\Lambda(A,B) = \{\mathrm{Id}\}\vee\Gamma$.

The matrix A is not multiconstrined and so has an adapted basis as above. We change the basis so that A and B are "canonical".

The proof is divided in cases according to the type of constraint. Except for the (4)-constrained case, the matrix A is diagonal, and so the space Γ is described by (2.6).

Unconstrained case: It follows from Lemma 2.13 that if A is unconstrained and diagonal then the only way for the pair (A,B) to be poor is that B has an off-diagonal zero entry. (The reader should review the proof of Lemma 2.13.)

(1)*-constrained case:* We see that the adjoint has two eigenvalues (different from 1) of multiplicity 2, namely $\lambda_1\lambda_2^{-1} = \lambda_2\lambda_3^{-1}$ and $\lambda_2\lambda_1^{-1} = \lambda_3\lambda_2^{-1}$. By the same reasoning as in the unconstrained case, it follows that $\{\mathrm{Id}\}\vee\Gamma$ contains the space

$$\big\{(y_{ij})\in\mathfrak{gl}(d,\mathbb{C});\ y_{11}=\cdots=y_{dd}\,,\ b_{12}^{-1}y_{12}=b_{23}^{-1}y_{23}\,,\ b_{21}^{-1}y_{21}=b_{32}^{-1}y_{32}\big\}.$$

This is a generalized Toeplitz space, and so by Example 2.2 it is transitive.

(2)*-constrained case:* The reasoning is very similar to that of the (1)-constrained case, but now the adjoint has four eigenvalues (different from 1) of multiplicity 2. The space $\Lambda(A,B)$ contains the following subspace:

$$\begin{aligned}\big\{(y_{ij})\in\mathfrak{gl}(d,\mathbb{C});\, &y_{11}=\cdots=y_{dd}\,,\ b_{13}^{-1}y_{13}=b_{24}^{-1}y_{24}\,,\\ &b_{12}^{-1}y_{12}=b_{34}^{-1}y_{34}\,,\ b_{21}^{-1}y_{21}=b_{43}^{-1}y_{43}\,,\ b_{31}^{-1}y_{31}=b_{34}^{-1}y_{34}\big\}.\end{aligned}$$

Again, this is a generalized Toeplitz space, and so it is transitive.

(3)*-constrained case:* This case is a little different from the two previous ones. The adjoint has an eigenvalue -1 of multiplicity 2. Recalling that b_{11} and b_{22} are different, and making use of the identity matrix, we see that $\Lambda(A,B)$ contains the following subspace:

$$\tilde{\Gamma} = \big\{(y_{ij})\in\mathfrak{gl}(d,\mathbb{C});\ y_{33}=\cdots=y_{dd}\,,\ b_{12}^{-1}y_{12}=b_{21}^{-1}y_{21}\big\}.$$

This is not a generalized Toeplitz space. However, consider the linear automorphism S that swaps the first two elements of the canonical basis of $\mathbb{C}^n$, and fixes the others. Then

$$S \cdot \tilde{\Gamma} = \left\{ (z_{ij}) \in \mathfrak{gl}(d, \mathbb{C});\ z_{33} = \cdots = z_{dd}\,,\ b_{12}^{-1} z_{22} = b_{21}^{-1} z_{11} \right\}$$

is a generalized Toeplitz space! By Remark 2.3, the space $S \cdot \tilde{\Gamma}$ is transitive, and so are $\tilde{\Gamma}$ and $\Lambda(A, B)$.

(4)-*constrained case:* This case is more involved because the operator Ad_A is not diagonalizable. We will explain its Jordan form. Let us explain visually how Ad_A acts: given any matrix, decompose it into blocks C_{ij} as in the following picture

$$\left(\begin{array}{c|c|c|c|c} C_{22} & C_{23} & C_{24} & \cdots & C_{2d} \\ \hline C_{32} & C_{33} & & \cdots & \\ \hline C_{42} & & C_{44} & \cdots & \\ \hline \vdots & \vdots & \vdots & \ddots & \\ \hline C_{d2} & & & \cdots & C_{dd} \end{array}\right)$$

where the block C_{22} is a 2×2 matrix, the blocks C_{2j} are 2×1, the blocks C_{i2} are 1×2 and the others are 1×1. Then, the operator Ad_A leaves invariant the space Γ_{ij} of matrices whose nonzero coefficients lie inside the block C_{ij}. Moreover, it is easily computed that the operator Ad_A has the following properties:

- restricting to the space Γ_{22}, which we canonically identify to $\mathfrak{gl}(2, \mathbb{C})$, one has:
$$\mathrm{Ad}_A \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}; \qquad \mathrm{Ad}_A \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$$
$$\mathrm{Ad}_A \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}; \qquad \mathrm{Ad}_A \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$
One then easily computes that, in the ordered basis formed by vectors
$$J_1 = \begin{pmatrix} 0 & -2 \\ 0 & 0 \end{pmatrix}; \quad J_2 = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}; \quad J_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}; \quad J_4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$
the matrix of $\mathrm{Ad}_A | \Gamma_{11}$ is
$$\left(\begin{array}{ccc|c} 1 & 1 & & \\ & 1 & 1 & \\ & & 1 & \\ \hline & & & 1 \end{array}\right).$$
- For any $j \geqslant 3$, identifying Γ_{2j} to the space of 2×1 matrices, the matrix of $\mathrm{Ad}_A | \Gamma_{2j}$ is $\begin{pmatrix} \lambda_2 \lambda_j^{-1} & 1 \\ 0 & \lambda_2 \lambda_j^{-1} \end{pmatrix}$ in the basis formed by matrices $\lambda_2 \lambda_j^{-1} E_{1,j} = \begin{pmatrix} \lambda_2 \lambda_j^{-1} \\ 0 \end{pmatrix}$ and $E_{2,j} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, where we use the notation $E_{i,j}$ from (2.4).
- For any $i \geqslant 3$, identifying $\Gamma_{i,2}$ to the space of 1×2 matrices, the matrix of $\mathrm{Ad}_A | \Gamma_{i2}$ is $\begin{pmatrix} \lambda_i \lambda_2^{-1} & 1 \\ 0 & \lambda_i \lambda_2^{-1} \end{pmatrix}$ in the basis formed by matrices $-\lambda_i \lambda_2^{-1} E_{i,1} = \begin{pmatrix} 0 & -\lambda_i \lambda_2^{-1} \end{pmatrix}$ and $E_{i2} = \begin{pmatrix} 1 & 0 \end{pmatrix}$.
- for $3 \leqslant i, j \leqslant d$, (E_{ij}) is a basis of Γ_{ij}; it is an eigenvector with eigenvalue $\lambda_i \lambda_j^{-1}$.

- The spaces Γ_{ij}, for $2 \leqslant i, j \leqslant d$ have respective spectra $\{\lambda_i \lambda_j^{-1}\}$, which for $i \neq j$ are pairwise disjoint and different from $\{1\}$.

The concatenation of the bases described above gives a Jordan basis for Ad_A. Now take a matrix B that is a good match for A, and consider its expression as a linear combination of the elements of that Jordan basis.

Claim A.4. *All coefficients in this linear combination are nonzero, except possibly the coefficients of the vectors J_1, J_2, J_4 and the vectors E_{ii}, for all $3 \leqslant i \leqslant d$.*

The verification is direct.

Consider now the splitting $\mathrm{Mat}_{d\times d}(\mathbb{C}) = V \oplus \Delta$, where Δ is the subspace $\mathbb{C}J_4 \oplus E_{33} \oplus \ldots \oplus E_{dd}$ of the space of diagonal matrices, and V is the space spanned by all other elements of the above Jordan basis. Note that

$$V = (\mathbb{C}J_1 + \mathbb{C}J_2 + \mathbb{C}J_3) \oplus \left(\bigoplus_{\substack{2\leqslant i,j\leqslant d \\ i\neq j}} \Gamma_{ij} \right)$$

is a decomposition of V into Ad_A-invariant subspaces with pairwise disjoint spectra. Let π be the projection onto V along Δ. It follows from the claim and Lemmas 3.1 and 3.2 that $\pi(B)$ is a cyclic vector for $\mathrm{Ad}_A|V$. So, using the Ad_A-invariance of the spaces V and Δ, we have

$$\pi(\Gamma) = \pi\big(\mathrm{sorb}_{\mathrm{Ad}_A}(B)\big) = \mathrm{sorb}_{\mathrm{Ad}_A}(\pi(B)) = V.$$

Note that V contains the matrices E_{ij}, for all $i \neq j$, hence $\{\mathrm{Id}\} \vee V$ is a generalized Toeplitz space. As π projects along a subspace of diagonal matrices, $\{\mathrm{Id}\} \vee \Gamma$ is again a generalized Toeplitz space and in particular is a transitive space.

We have considered the four types, and Proposition A.2 is proved. $\square$

Appendix B. Complementary facts about singular constant inputs of generic type

In this appendix we give grounds for Remark 1.3. We also discuss other control-theoretic properties of generic semilinear systems, related to universal regularity.

B.1. **Local persistence of singular inputs.** Let $A \in C^r(\mathcal{U}, \mathrm{GL}(d,\mathbb{R}))$, $r \geqslant 1$. We will work upon Lemma 2.9 in order to obtain a more practical way to detect that the 1-jet of A at a point corresponds to conspicuously poor data. (Recall from Remark 6.7 that this is the only type of poor data that appears generically.) For example, in the $m = 1$, $d = 2$ case, we will see that conspicuous poorness means that the angular velocity of one of the eigendirections vanishes (see Remark B.1 below).

Suppose that $u_0 \in \mathcal{U}$ is such that the matrix $A(u_0)$ is diagonalizable over $\mathbb{R}$ and with simple eigenvalues only. By Proposition 2.10, there is a neighborhood $\mathcal{U}_0$ of u_0 and C^r-maps $\lambda_1, \ldots, \lambda_d \colon \mathcal{U}_0 \to \mathbb{C}$ such that for all $u \in \mathcal{U}_0$, the complex numbers $\lambda_i(u)$ are all distinct, and form the spectrum of $A(u)$; moreover there exist a C^r map $P \colon \mathcal{U}_0 \to \mathrm{GL}(d,\mathbb{R})$ such that for all $u \in \mathcal{U}_0$,

$$A(u) = P(u)\,\Delta(u)\,P^{-1}(u)\ , \text{ where } \Delta(u) = \mathrm{Diag}(\lambda_1(u), \ldots, \lambda_d(u)). \tag{B.1}$$

For simplicity, let us consider first case where $\mathcal{U}$ is an interval in $\mathbb{R}$ (in particular $m = 1$). Then the normalized derivative of A at a point u can be identified with $N(u) := A'(u)\,A^{-1}(u)$. Consider the expression of $N(u)$ in the basis that diagonalizes $A(u)$, that is, $B(u) := P^{-1}(u)\,N(u)\,P(u)$. Since $\frac{\mathrm{d}}{\mathrm{d}u}P^{-1}(u) = -P^{-1}(u)\,P'(u)\,P^{-1}(u)$, we compute that

$$B(u) = \Delta'(u)\,\Delta^{-1}(u) + Q(u) - \Delta(u)\,Q(u)\,\Delta^{-1}(u)\,,$$

where

$$Q(u) := P^{-1}(u)\, P'(u)\,.$$

So the off-diagonal entries of the matrices $B(u)$ and $Q(u)$ are related by

$$b_{ij}(u) = \big(1 - \lambda_i(u)/\lambda_j(u)\big)\, q_{ij}(u) \quad (i \neq j).$$

In view of Lemma 2.9, we conclude the following: if for some $u_* \in \mathcal{U}_0$

(B.2) there is an off-diagonal entry position (i,j) such that $q_{ij}(u_*) = 0$

then the 1-jet $j^1 A(u_*)$ is poor.

Remark B.1. Let us give a geometrical interpretation of condition (B.2). The columns of P form a basis $(v_1, \dots, v_d)$ of eigenvectors of A, and the rows of P^{-1} form a basis $(f_1, \dots, f_d)$ of eigenfunctionals of A (in the sense that $f_i \circ A = \lambda_i f_i$); these two bases are related by $f_i(v_j) = \delta_{ij}$. So $q_{ij} = f_i\left(\frac{\mathrm{d}v_j}{\mathrm{d}u}\right)$ is the component of the velocity of v_j in the direction of v_i. For example, for $d = 2$, condition (B.2) means that one of the eigendirections of A has zero angular speed at instant $u = u_*$.

It is trivial to adapt the previous calculations to the higher dimensional case and then conclude the following:

Proposition B.2. *Let $(u_1, \dots, u_m)$ be coordinates in a chart domain $\mathcal{U}_0 \subset \mathcal{U}$ where expression* (B.1) *holds. Consider matrices*

$$Q_k(u) := P^{-1}(u)\, \frac{\partial P}{\partial u_k}(u)\,. \tag{B.3}$$

If for some $u_* \in \mathcal{U}_0$ there is an off-diagonal entry position (i,j) such that

(B.4) *for each $k = 1, \dots, m$, the (i,j)-entry of the matrix $Q_k(u_*)$ vanishes*

then the 1*-jet $j^1 A(u_*)$ is poor, that is, the constant input $(u_*, \dots, u_*)$ (of any length) is singular.*

In the situation of Proposition B.2, assume additionally that the map

$$\Phi\colon \begin{cases} \mathcal{U}_0 & \to \operatorname{Im}\Phi \subset \mathbb{K}^m \\ u & \mapsto [\text{the } (i,j)\text{-entry of } Q_k(u)]_{1 \leqslant k \leqslant m} \end{cases} \quad \text{is a diffeomorphism.} \tag{B.5}$$

In that case, the existence of a poor jet is persistent in the following way: If $\tilde{A}$ is sufficiently C^2-close to A then by Proposition 2.10 we can express $\tilde{A}(u) = \tilde{P}(u)\, \tilde{\Delta}(u)\, \tilde{P}^{-1}(u)$ for u close to u_*, where $\tilde{P}$ and $\tilde{\Delta}$ are C^2-close to P and Δ respectively, and $\tilde{\Delta}$ is diagonal. The corresponding matrices $\tilde{Q}_k = \tilde{P}^{-1} \frac{\partial \tilde{P}}{\partial u_k}$ are C^1-close to Q_k and the map

$$\tilde{\Phi}\colon u \mapsto \Big[\text{the } (i,j)\text{-entry of } \tilde{Q}_k(u)\Big]_{1 \leqslant k \leqslant m}$$

is C^1-close to Φ. By (B.5) the fact that $\Phi(u_*) = (0, ..., 0)$, there is $\tilde{u}$ close to u_* such that $\tilde{\Phi}(u) = (0, ..., 0)$. In particular the 1-jet $j^1 \tilde{A}(\tilde{u})$ is poor.

Now, concerning existence: It is evident that a domain $\mathcal{U}_0$ and 2-jets $j^2 P(u_*)$ satisfying conditions (B.4) and (B.5) actually exist; moreover we can always find a map $P\colon \mathcal{U} \to \mathrm{GL}(d, \mathbb{R})$ with a prescribed 2-jet at a point u_*. In view of the discussion above, we conclude the following:

Proposition B.3 (Persistence of singular inputs)**.** *For any $d \geqslant 1$ and any d-dimensional smooth manifold $\mathcal{U}$, there exists a C^2-open nonempty subset of maps $A \in C^2(\mathcal{U}, \mathrm{GL}(d, \mathbb{R}))$ such that the following holds:*

there exists $u \in \mathcal{U}$ such that the constant inputs $(u, \dots, u)$ of any length are all singular for the system (1.4)*.*

That is, one cannot improve Theorem 1.1 replacing "discrete set" by "empty set".

We can also see why the statement of Theorem 1.1 with "C^2-open" replaced by "C^1-open" is not true: Given any map A such that (B.4) holds at some point, we can C^1-perturb A (by C^0-perturbing P) in a way such that (B.4) now holds for a non-discrete set of points.[18]

B.2. **Other control-theoretic properties.** We now introduce a few control-theoretic notions related to accessibility and regularity, and discuss the validity of statements similar to Theorem 1.1 for these notions.

Consider a general control system (1.1). Fix a time length N, and let ϕ_N denote the response map as in (1.2). We say that a trajectory determined by $(x_0; u_0, \dots, u_{N-1})$ is:

- *locally accessible*[19] if for every neighborhood V of $(u_0, \dots, u_{N-1})$ in $\mathcal{U}^N$, the set $\phi_N(\{x_0\} \times V)$ has nonempty interior.
- *strongly locally accessible* if for every neighborhood V of $(u_0, \dots, u_{N-1})$ in $\mathcal{U}^N$, the set $\phi_N(\{x_0\} \times V)$ contains in its interior the final state $\phi_N(x_0; u_0, \dots, u_{N-1})$.

The following implications are immediate:

$$\text{regular} \Rightarrow \text{strongly locally accessible} \Rightarrow \text{locally accessible}.$$

We say that an input $(u_0, \dots, u_{N-1})$ is *universally locally accessible* (resp. *universally strongly locally accessible*) if the trajectory determined by $(x_0; u_0, \dots, u_{N-1})$ is locally accessible (resp. strongly locally accessible).

Now we come back to the context of projective semilinear control systems (1.4). A (relatively weak) corollary of Theorem 1.1 is that for generic maps A, universal local accessibility holds at all constant inputs:

Proposition B.4. *Let $N \in \mathbb{N}$ and $\mathcal{O} \subset C^2(\mathcal{U}, \mathrm{GL}(d, \mathbb{R}))$ be as in Theorem 1.1. For any $A \in \mathcal{O}$, every constant input sequence of length N is universally locally accessible.*

Proof. If $A \in \mathcal{O}$ then for every constant input sequence of length N we can find a regular input sequence nearby. □

As we have shown in Proposition B.3, it is not possible to improve Proposition B.4 by replacing "local accessibility" by "regularity". Neither it is possible to replace "local accessibility" by "strong local accessibility", as the following simple example (in $m = 1$, $d = 2$) shows:

Example B.5. For $u \in \mathbb{R}$, define

$$P(u) = \begin{pmatrix} 1 & u \\ u^2 & 1 \end{pmatrix}, \quad \Delta(u) = \mathrm{Diag}(2, 1).$$

Let $\mathcal{U}$ be an small open interval containing 0, and define $A \colon \mathcal{U} \to \mathrm{GL}(2, \mathbb{R})$ by (B.1). Let $\xi_0 \in \mathbb{RP}^1$ correspond to the direction of the vector $(1, 0)$. Then for any subinterval $V \ni 0$, and any $N > 0$, the set

$$\phi_N(\{\xi_0\} \times V^N) = \{A(u_{n-1}) \cdots A(u_0) \cdot \xi_0 \; u_i \in V\}$$

is an "interval" of $\mathbb{RP}^1$ containing $\xi_0 = \phi_N(\xi_0; 0, \dots, 0)$ in its boundary. Therefore the input $(0, \dots, 0)$ is not universally strongly locally accessible. A similar situation occurs for any C^2-perturbation of A.

[18] Using this idea and Baire's theorem, one can also show that the conclusion of Theorem 1.1 is not true for C^1-generic maps A; actually for C^1-generic A, the points $u \in \mathcal{U}$ corresponding to singular constant controls form a perfect set.

[19] Beware: a different concept with this name appears in [CK2].

Appendix C. Proof of a complex version of Theorem 1.1

For the complex case we consider instead holomorphic mappings $A : \mathcal{U} \to \mathrm{GL}(d, \mathbb{C})$.

More precisely, given an open subset $\mathcal{U} \subset \mathbb{C}^m$, we denote by $\mathcal{H}(\mathcal{U}, \mathrm{GL}(d, \mathbb{C}))$ the set of holomorphic mappings $A\colon \mathcal{U} \to \mathrm{GL}(d, \mathbb{C})$ endowed with the usual topology of uniform convergence on compact sets.

Theorem C.1. *Given integers $d \geqslant 2$ and $m \geqslant 1$, there exists an integer $N \geqslant 1$ with the following properties. Let $\mathcal{U} \subset \mathbb{C}^m$ be an open subset. Then for any compact set $K \subset \mathcal{U}$, there exists an open and dense subset $\mathcal{O}$ of $\mathcal{H}(\mathcal{U}, \mathrm{GL}(d, \mathbb{C}))$ such that for any $A \in \mathcal{O}$ the constant inputs in $\mathbb{K}^N$ are all universally regular for the system* (1.4)*, except for a finite subset.*

We have the straightforward corollary:

Corollary C.2. *Given integers $d \geqslant 2$ and $m \geqslant 1$, there exists an integer $N \geqslant 1$ with the following properties. Let $\mathcal{U} \subset \mathbb{C}^m$ be an open subset. There exists a residual subset $\mathcal{R}$ of $\mathcal{H}(\mathcal{U}, \mathrm{GL}(d, \mathbb{C}))$ such that for any $A \in \mathcal{R}$ the constant inputs in $\mathcal{U}^N$ are all universally regular for the system* (1.4)*, except for a discrete subset.*

These results could probably be obtained in certain more general complex manifolds. But in order to avoid technicalities, we consider only open subsets of $\mathbb{C}^m$. Also, we use only elementary real transversality tools.

Proof of Theorem C.1. Let $\mathcal{U} \subset \mathbb{C}^m$ be an open subset. We may identify the set of 1-jets from $\mathcal{U}$ to $\mathrm{GL}(d, \mathbb{C})$ with

$$\mathcal{U} \times GL(d, \mathbb{C}) \times \mathfrak{gl}(d, \mathbb{C})^m.$$

As we did in Section 6, and using Theorem 1.9 instead of Theorem 1.8, we obtain that the set of poor 1-jets from $\mathcal{U}$ to $\mathrm{GL}(d, \mathbb{C})$ is the algebraic subset $\mathcal{U} \times \mathcal{P}_m^{(\mathbb{C})}$ of the space of 1-jets. Hence it admits a stratification

$$\mathcal{U} \times \mathcal{P}_m^{(\mathbb{C})} = \mathcal{U} \times \Sigma_n \supset \ldots \supset \mathcal{U} \times \Sigma_0.$$

Write $\mathcal{U} \times \mathcal{P}_m^{(\mathbb{C})}$ as the disjoint union $\bigsqcup_{0 \leqslant i \leqslant n} X_i$ where each X_i is a smooth submanifold of dimension i in the jet space $J^1(\mathcal{U}, \mathrm{GL}(d, \mathbb{C}))$, and X_n has codimension m.

Fix now a map $A \in \mathcal{H}(\mathcal{U}, \mathrm{GL}(d, \mathbb{C}))$. For all $v = (a, b_1, \ldots b_m) \in \mathbb{C}^{m+1}$ and $u = (u_1, \ldots, u_m) \in \mathbb{C}^m$, write

$$P_v(u) = a + \sum_{i=1}^{m} b_k u_k.$$

For all $\mathtt{v} = (v_{i,j})_{1 \leqslant i,j \leqslant d} \in \left(\mathbb{C}^{m+1}\right)^{d^2}$, write $P_{\mathtt{v}} = \left[P_{v_{i,j}}\right]_{1 \leqslant i,j \leqslant d}$ and define the map $\Phi_{\mathtt{v}} = A + P_{\mathtt{v}}$. One can write the 1-jet extension $j^1 A$ at the point $u \in \mathcal{U}$ as

$$j^1 A(u) = [u, A(u), B_1, \ldots, B_m] \in \mathcal{U} \times \mathrm{GL}(d, \mathbb{C}) \times [\mathrm{Mat}_{d \times d}(\mathbb{C})]^m.$$

The same way, if we put $v_{i,j} = (a_{i,j}, b_{1,i,j}, \ldots, b_{m,i,j})$, we have

$$j^1 P_{\mathtt{v}}(u) = [u, P_{\mathtt{v}}(u), (b_{1,i,j})_{1 \leqslant i,j \leqslant d}, \ldots, (b_{m,i,j})_{1 \leqslant i,j \leqslant d}].$$

Define the map $F\colon v \mapsto F_{\mathtt{v}} = j^1 \Phi_{\mathtt{v}}$. The evaluation map of F is:

$$F^{\mathrm{ev}}\colon \begin{cases} \left(\mathbb{C}^{m+1}\right)^{d^2} \times \mathcal{U} & \to \ \mathcal{U} \times \mathrm{Mat}_{d \times d}(\mathbb{C}) \times [\mathrm{Mat}_{d \times d}(\mathbb{C})]^m \\ (\mathtt{v}, u) & \mapsto \ F_{\mathtt{v}}(u) \end{cases}.$$

Hence,

$$\begin{aligned}F^{\mathrm{ev}}(\mathtt{v}, u) &= j^1(A + P_{\mathtt{v}}) \\ &= \Big[u, (A + P_{\mathtt{v}})\,(u), (b_{1,i,j})_{1\leqslant i,j\leqslant d}\,, \ldots, (b_{m,i,j})_{1\leqslant i,j\leqslant d}\Big]\end{aligned}$$

Claim C.3. *For all u, the map F^{ev} restricts to a submersion from the $(\cdot, u)$-fiber to the $[u, \cdot]$-fiber.*

Proof. We want to prove that

$$\mathtt{v} \mapsto \Big[(A + P_{\mathtt{v}})(u), (b_{1,i,j})_{1\leqslant i,j\leqslant d}\,, \ldots, (b_{m,i,j})_{1\leqslant i,j\leqslant d}\Big]$$

is a submersion, or equivalently that

$$\mathtt{v} \mapsto \Big[P_{\mathtt{v}}(u), (b_{1,i,j})_{1\leqslant i,j\leqslant d}\,, \ldots, (b_{m,i,j})_{1\leqslant i,j\leqslant d}\Big]$$

is a submersion. Noting that $\mathtt{v} = (a_{i,j}, b_{k,i,j})_{\substack{1\leqslant i,j\leqslant d \\ 1\leqslant k\leqslant m}}$, this comes easily from the fact that $(a_{i,j}) \mapsto P_{\mathtt{v}}(u)$ is a submersion, for any fixed set of coefficients $(b_{k,i,j})_{\substack{1\leqslant i,j\leqslant d \\ 1\leqslant k\leqslant m}}$. ☐

That claim immediately implies that F^{ev} is a submersion. In particular it is transverse to each X_i. By the parametric transversality theorem (see [Hi, p. 79]), there is a residual subset of parameters $\mathtt{v}$ in $\left(\mathbb{C}^{m+1}\right)^{d^2}$ such that $F_{\mathtt{v}} = j^1\Phi_{\mathtt{v}}$ is transverse to X_i, for all i.

When $\mathtt{v}$ goes to 0, $\Phi_{\mathtt{v}}$ tends to A in $\mathcal{H}(\mathcal{U}, \mathrm{GL}(d, \mathbb{C}))$. Hence, the denseness in $\mathcal{H}(\mathcal{U}, \mathrm{GL}(d, \mathbb{C}))$ of the maps $\hat{A}$ such that $j^1\hat{A}$ is transverse to X_i, for all i. Take such a map $\hat{A}$: for all i, the image of $j^1\hat{A}$ does not intersect $X_0 \sqcup \cdots \sqcup X_{n-1}$ and intersects X_n (which has codimension m) only in a discrete subset.

Fix $K' \subset \mathcal{U}$ a compact set that contains K in its interior. The image $j^1\hat{A}$ restricted to K' can only intersect X_n in a finite set Γ: indeed, any accumulation point of that intersection set would have to be in $X_0 \sqcup \cdots \sqcup X_{n-1}$, since $X_0 \sqcup \ldots \sqcup X_n$ is closed, and this would contradict that $j^1\hat{A}$ does not intersect $X_0 \sqcup \cdots \sqcup X_{n-1}$.

By the choice of our topology, a small perturbation $\tilde{A}$ of $\hat{A}$ is C^0 close to $\hat{A}$ by restriction to K'. By Cauchy's formula, the map $\tilde{A}$ is C^2 close to $\hat{A}$ over the set K. Hence, the (compact) image of $j^1\tilde{A}$ restricted to K is still far from $X_0 \sqcup \cdots \sqcup X_{n-1}$, and intersects X_n transversally in some ϵ-neighborhood of Γ inside X_n. Thus it also has to intersect X_n in only a finite set.

So we found an open and dense subset of holomorphic maps whose 1-jets above K intersect the set of N-poor jets in only on a finite number of points. As a consequence, for such maps, there are only a finite number constant singular inputs in K^N for the system 1.4. This concludes the proof of Theorem C.1. ☐

Acknowledgements. We are grateful for the hospitality of Institute Mittag–Leffler, where this work begun to take form in 2010. We thank R. Potrie, L. San Martin, S. Tikhomirov, and C. Tomei for valuable discussions.

References

[Ar] Arnold, V.I. *Geometrical methods in the theory of ordinary differential equations.* 2nd edition. Springer, 1988. (Cited on page 24.)

[Az] Azoff, E.A. On finite rank operators and preannihilators. *Mem. Amer. Math. Soc.* 64, no. 357 (1986). (Cited on pages 7 and 14.)

[BG1] Bochi, J.; Gourmelon, N. Note on the dimension of certain algebraic sets of matrices. arXiv:1201.2217 (Cited on pages 6 and 30.)

[BG2] ———. Transitivity of spaces of matrices. In preparation. (Cited on pages 8 and 30.)

[BG3] ______. Universal regular control for generic semilinear systems: general inputs. In preparation. (Cited on pages 3 and 6.)

[BCR] BOCHNAK, J.; COSTE, M.; ROY, M.-F. *Real algebraic geometry.* Springer, 1998. (Cited on pages 8 and 30.)

[CK1] COLONIUS, F.; KLIEMANN, W. Linear control semigroups acting on projective space. *J. Dynam. Differential Equations* 5 (1993), no. 3, 495–528. (Cited on page 2.)

[CK2] ______. The dynamics of control. Birkhäuser, 2000. (Cited on pages 2 and 44.)

[El] ELLIOTT, D.L. *Bilinear control systems.* Springer, 2009. (Cited on page 4.)

[Ga] GANTMACHER, F.R. *The theory of matrices. Vol. 1.* Chelsea Publ. Co., 1959. Reprinted by the AMS, 2000. (Cited on pages 12, 13, and 24.)

[GWPL] GIBSON, C.G.; WIRTHMÜLLER, K.; DU PLESSIS, A.A.; LOOIJENGA, E.J.N. *Topological stability of smooth mappings.* Lecture Notes in Mathematics, Vol. 552. Springer, 1976. (Cited on page 35.)

[Ha] HARRIS, J. *Algebraic geometry: a first course.* Springer, 1992. (Cited on pages 29 and 30.)

[Hi] HIRSCH, M.W. *Differential Topology.* Springer, 1976. (Cited on pages 2, 36, and 46.)

[Ka] KATO, T. *Perturbation theory for linear operators.* Reprint of the 1980 edition. Springer, 1995. (Cited on page 10.)

[La] LANG, S. *Algebra.* Revised 3rd edition. Springer, 2002. (Cited on page 34.)

[Le] LEVI, F.W. Ordered groups. *Proc. Indian Acad. Sci.* 16 (1942), 256–263. (Cited on page 16.)

[Ma] MATHER, J.N. Stratifications and mappings. *Dynamical systems (Proc. Sympos., Univ. Bahia, Salvador, 1971)*, 195–232. Academic Press, New York, 1973. (Cited on page 35.)

[Sh] SHAFAREVICH, I.G. *Basic Algebraic Geometry. Vol 1.* 2nd edition. Springer, 1994. (Cited on page 29.)

[So] SONTAG, E.D. Universal nonsingular controls. *Systems Control Lett.* 19 (1992), no. 3, 221–224. Errata: Ibid, 20 (1993), no. 1, 77. (Cited on page 2.)

[SW] SONTAG, E.D.; WIRTH. F.R. Remarks on universal nonsingular controls for discrete-time systems. *Systems Control Lett.* 33 (1998), no. 2, 81–88. (Cited on pages 2 and 29.)

[Tr] TROTMAN, D.J.A. Stability of transversality to a stratification implies Whitney (a)-regularity. *Invent. Math.* 50 (1978/79), no. 3, 273–277. (Cited on page 36.)

[Wi] WIRTH, F. Dynamics of time-varying discrete-time linear systems: spectral theory and the projected system. *SIAM J. Control Optim.* 36 (1998), no. 2, 447–487. (Cited on page 2.)

[Wh] WHITNEY, H. Elementary structure of real algebraic varieties. *Ann. of Math.* 66 (1957), 545–556. (Cited on page 31.)

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO (PUC–RIO)
URL: `www.mat.puc-rio.br/~jairo`
E-mail address: `jairo@mat.puc-rio.br`

INSTITUT DE MATHÉMATIQUES DE BORDEAUX, UNIVERSITÉ BORDEAUX I
E-mail address: `nikolaz.gourmelon@gmail.com`